Presented to
UNITED
Bonebrake
Theological Seminary
Library
Dayton, Ohio

by

William L. Sanders

TESTING INTELLIGENCE AND ACHIEVEMENT

THE MACMILLAN COMPANY
NEW YORK · BOSTON · CHICAGO · DALLAS
ATLANTA · SAN FRANCISCO

MACMILLAN & CO., LIMITED
LONDON · BOMBAY · CALCUTTA
MELBOURNE

THE MACMILLAN CO. OF CANADA, LTD.
TORONTO

TESTING INTELLIGENCE AND ACHIEVEMENT

BY

ALBERT J. LEVINE, PH.D.

PUBLIC SCHOOLS, NEW YORK CITY

AND

LOUIS MARKS, M.A.

EXAMINER, BOARD OF EDUCATION, NEW YORK CITY
CO-AUTHOR OF "FITTING THE SCHOOL TO THE CHILD"

LIBRARY OF
United Theological Seminary
BF
431
.L4
32,020

New York

THE MACMILLAN COMPANY

1928

All rights reserved

Copyright, 1928,
By THE MACMILLAN COMPANY.

Set up and electrotyped. Published November, 1928.

PREFACE

This book is intended to serve teachers, social workers, and students of psychology as a compact and rapid survey of the field of testing. It is designed for use as a textbook, and combines the brevity of the outline with the thoroughness of the monograph. The authors have striven to give the facts of experimental psychology in a context concise enough to serve as a ready reference and expansive enough to aid in the easy recall of the elementary principles which lead to their formulation. Whenever possible, the material has, therefore, been arranged in the form of quickly assimilable outlines and has been organized under headings and subheadings, topics and subtopics, so that memorization is greatly facilitated.

The authors gratefully acknowledge the substantial aid extended ungrudgingly by Dr. Paul Klapper, Dean of the College of the City of New York, and by Dr. Egbert Turner of the Faculty of Education. No less valuable has been the whole-hearted coöperation of James A. O'Donnel, since deceased, Principal of Isaac Remsen Junior High School, during the preparation of the chapters on individual and group tests.

A. J. L.
L. A. M.

UNITED THEOLOGICAL SEMINARY
AUG 5 1954

CONTENTS

		PAGE
PREFACE		v
CHAPTER		
I.	THE TESTING MOVEMENT	1
	Salient Phases	3
	Summary	32
II.	AN ANALYSIS OF PERSONALITY	38
	A. Innate Equipment	40
	B. Acquired Equipment	56
	C. Directive Equipment	83
III.	INTELLIGENCE: WHAT IT IS AND HOW IT IS MEASURED	96
	A. What Intelligence Is	96
	B. How Intelligence Is Tested	107
	C. Educational Measures of Intelligence	127
IV.	REPRESENTATIVE INDIVIDUAL SCALES AND TESTS	137
	A. Principal Scales of Intelligence	137
	B. Supplementary Scales of Intelligence	152
V.	GROUP TESTS OF INTELLIGENCE	161
	Representative Group Tests	162
VI.	ACHIEVEMENT TESTS AND SCALES	168
	A. As Objective Measurements	188
	B. Tests and Scales	192
VII.	TESTING NON-INTELLIGENCE TRAITS OF PERSONALITY	238
	A. The Educational Significance of Non-Intelligence Personality Traits	241
	B. Studies in the Nature of Character	266
	C. Summary and Conclusion	286

CHAPTER PAGE

VIII. The Mental Defective 306
A. Historical Résumé 306
B. Characteristics of the Mental Defective . . 310
C. Testing and Establishing Mental Deficiency . 314
D. Teaching the Mental Defective . . . 319

IX. The Neurotic 328
A. Difference between the Defective and the Neurotic 328
B. Characteristics of the Psychopatic Individual . 329
C. Causes of Psychopathy 330
D. Means of Determining Psychopathy . . . 338
E. The School and the Neurotic 343

X. The Superior Child 347
A. Attitude toward the Superior Child . . . 347
B. Characteristics of the Superior Child . . 349
C. Education of the Superior Child . . . 359

Final Test 369
Summary of Statistical Symbols, Terms, and Formulæ in Educational Measurements . 383
Index 393

TESTING INTELLIGENCE AND ACHIEVEMENT

TESTING INTELLIGENCE AND ACHIEVEMENT

CHAPTER I

THE TESTING MOVEMENT

Mankind has a passion for both the abstract and the concrete. It is just as eager to plumb the depths of the soul as it is avid to measure accurately a hair's breadth. It can evolve the most intricate of philosophies or the most delicate of scales. Sometimes it happens that the library and the laboratory become intimate explorers in the same realm of knowledge. The history of the testing movement is woven out of the strands of philosophy, metaphysics, psychology, physics, and mathematics. The fabric lacks evenness of texture, coarseness and fineness alternating throughout its length. This is to be expected when we consider the extreme mobility of mind and the rigid machinery of arithmetic. You may tap the consciousness and draw off a thought; you can never correctly measure it, for the moment it leaves the main stream it is already a broken thought, without support and without sustenance.

Yet some sort of quantitative organization of psychological data has been successfully attempted. We call it generically mental testing, because of the overwhelming importance of intellect in psychology. In the school, in the shop, in the street, it was the quality of mind that insured success. The feelings and emotions were supposed to occupy distinct compartments. One learned with his brain in school and in the shop, and one got along with his neighbor by virtue of good brain power. The dull boy and the bad boy were both indiscriminately classed as "stupid"; the dunce cap was the universal index of a poor mind. Again, behavior of an overwhelmingly emotional source was also traced to the brain; the idiot was "out of his mind." Not that feeling and willing did not exist as psychological categories; but both laymen and psychologists seem to have fallen into one of two attitudes: either they regarded the mind as an expression of knowing, willing, and feeling, conceiving it to be the normal channel for their discharge, in which they were restrained as a river by its banks; or they thought knowing, feeling, and willing distinct entities, functioning independently of each other. Knowing, however, was accorded greater deference, for its direct intervention produced concrete, definite, and calculable results. Mind became synonymous with achievement; feeling and willing may have made vital contributions, but the mind evolved the strategy.

You will recognize in both attitudes the old "faculty"

psychology. It is slowly yielding before the advance of the new psychology, but it is still with us in the guise of "categories." After all, the human mind is a complicated machine, and behavior is no less subtle in its varied manifestations. The unitary conception of mind does not always find a logical place for many a mobile tendril of personality that escapes beyond the boundaries of the concept.

Salient Phases

The measuring movement, like all educational movements, cannot boast of continuous development. It began in insignificant undertakings at isolated places — small springs of knowledge that either dried up or overflowed in a gentle current forming in time a definite channel. Other channels soon sprang into being, received in turn small tributaries, and in time effected a junction. The new current gained in volume and momentum. A movement was born.

It is our purpose to emphasize the fresh accessions to volume rather than merely to trace its course. Chronology will not enter greatly into our discussion because there is a deplorable lack of fixity in dates connected with mental testing. Again, many educational experiments have been first described in magazine articles. The historian is sometimes hard put to it to decide whether the formulation of a certain principle shall be dated from the time it appeared in magazine or in book form. Nor is this the only difficulty that faces him.

It is becoming increasingly evident that education is international in scope; that contributions to mental measurement come from many lands; and that endless research awaits the historian who would bring to light the masses of experimental data languishing in forgotten magazines or in unpublished theses. Even American material still awaits the student. It is largely owing to the fragmentary character of our knowledge that so many educators are prone to hail with acclaim the publication of an investigation whose findings have been already discredited, but of whose existence the investigator and reader have remained in ignorance. The contradictory nature of so many of the articles on mental testing is the despair of the thoughtful student. Subordinating, then, the chronological element, we shall pursue the testing movement through four of its principal phases, bearing in mind that crystallization is still in the distant future and that progress entails both advances and retreats. The four phases are:

1. Social qualitative
2. Psychological qualitative
3. Psychological quantitative
4. Biological qualitative

(1) *Social Qualitative Phase*

The intelligence test is not a modern institution. The Greeks used a rough measure of efficiency in condemning undesirable infants to exposure. The feeble-minded were eliminated as well as the physically

defective. The purpose was social — the betterment of the race. Mental defect uncomplicated by physical deterioration was not considered objectionable. Among the Romans exposure was not so common.

Christianity dealt much more humanely with the mentally defective. However, the attitude of sympathy and pity soon hardened into cruel intolerance. The idiot was first tolerated, then reverenced, then burnt, according as charity or fanaticism ruled. The era of monastic asylum was succeeded by "the era of whips and chains."

In England the insistence for segregating defectives led to the establishment of norms for measuring sanity. In Sir Anthony Fitzherbert's *The New Natura Brevium* (1534) we find this reference: "And he who shall be said to be a sot and idiot from his birth is such a person who cannot account or number twenty pence, nor can tell who was his father and mother, nor how old he is, etc., so it may appear that he hath no understanding of reason what shall be for his profit, nor what for his loss."

But the ravages of defectiveness continued unabated, and society became acutely aware of the inadequacy of this qualitative test. The line between sanity and insanity, normality and abnormality, remained but dimly demarcated. In 1797 occurred an event of great importance. A wild boy was found in the woods of Cannes, Department of Aveyron, France. He was brought to Paris, where he was placed in charge of

Itard, the physician of the Institution for Deaf-Mutes. Some notable advances had already been made in teaching the blind and the deaf-mute, so that some hardy adventurer here and there dared to flout the medical dictum that idiots, being "human brutes," were incurable and uneducable. Itard therefore endeavored to educate this boy, but gave up the task in disgust.

His pupil Seguin saw that the training had brought about an improvement, and he set out to test the theory of the unimprovability of idiots and imbeciles. He later set down his ideas in *Traitment Moral, Hygiene, et Education des Idiots.* Seguin arrived at many remarkable conclusions, clarifying the views of the advanced thinkers with regard to the true inwardness of feeble-mindedness. His ideas gained greater decisiveness in another book, which he wrote in the United States (1864). It is called *Idiocy: Its Diagnosis and Treatment by the Physiological Method.*

(2) *Psychological Qualitative Phase*

It may be said that the psychological qualitative phase of the testing movement began with Seguin's later book. Heretofore idiocy was the exclusive concern of legislators and physicians, who saw in this affliction a purely sociological and medical problem. The medical profession was aware of the mental aspects of the disease, but they studied them as symptoms and not as causes. Seguin, though a physician, was led

through teaching to the consideration of the purely psychological factors, which he came to view as the causal factors in feeble-mindedness. His psychology sounds truly modern in that he does not decentralize the mind from the rest of the personality. To this extent his work is biologic in its trend. "Man," he decides, "being a unit, is artificially analyzed for study into activity, intelligence, and will. . . . The predominance of any one of these functions constitutes a disease, their perversion leads to insanity, their notable deficiency at birth constitutes idiocy, afterwards imbecility, later dementia."

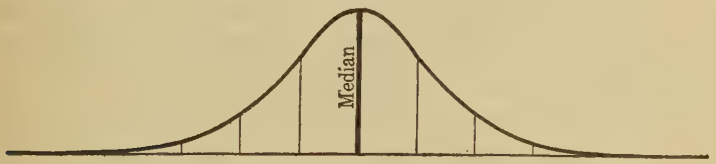
Seguin and his followers continued to study ways of delineating sharply the insane from the sane, and the idiots from the imbeciles, on one hand, and the feeble-minded, on the other. Such descriptive terms as "abnormal," "backward," "idiot," "mentally deficient," "unstable" soon made their appearance, testifying to the growing interest in abnormalities. As classificatory terms they were inexact and brought in their train the evils of disparate language. . Their lack of preciseness was keenly felt, but without tests, conditions could but tend to multiplication of terminology. Some advance, however, was made. Equirol made the important distinction between the dement and the idiot, and Itard, Fabret, and Voisin attempted to classify the symptoms of idiocy. Psychology had supplied a workable theory ; the measuring rod, however, was still subjective-qualitative.

(3) *Psychological Quantitative Phase*

The psychological quantitative phase began definitely with Francis Galton (1822–1911). It first defined itself as early as 1897. In that year Wundt opened his laboratory in Leipsic. Experimental psychology was born and grew into lusty adulthood under the guidance of Fechner and Titchener. The impulse they gave to quantitative evaluation of psychological processes had its repercussion in other fields, and psychological laboratories soon hummed with activity directed toward the measurement of sensations and images. The aim was purely scientific and did not concern itself with the applied psychology of mental testing. As Titchener phrases it, "Science goes on its way without regard to human interests and without aiming at any practical goal."

Soon psychological data multiplied: fatigue, discrimination of brightness, dual cutaneous impressions, and rote memory. Francis Galton carried the scientific method into the field of associative learning and individual differences. He imported into his studies a more exact technique applying, for the first time, the mathematic method of Gauss and Quetelet. He argued that all social phenomena range themselves in a curve (bell-shaped) of probability, which beginning at one end with subaverage measurements rapidly increase in number so that the overwhelmingly large class of average scores range themselves solidly around the

perimeter of the bell and gradually slope down to include the relatively smaller number of the superior. He also applied to these data the method of averages and dispersion, measures indicating the gradations from lowest and highest scores to the average. His mathematical zeal was actuated by the ideal "that until the phenomena of any branch of knowledge have been subjected to measurement and number, it cannot assume the status and dignity of a science."



NORMAL CURVE OF DISTRIBUTION

Galton's influence upon the mental testing movement is enormous. His first edition of *Inquiries into Human Faculty* contained studies in imagery, traits, and statistical methods. In *Hereditary Genius* he turned to heredity, thus enriching the subject of mental testing by opening a new field for investigation. We are still beset by torturing doubts as to the heritability of mental endowment and the constancy of the "intelligence quotient." During the 'eighties, Galton established a eugenics laboratory at the South Kensington Museum in London, where one might have his physical and mental capacities measured.

Side by side with refinement in methodology may be placed the work of the army of investigators, who,

fired by the brilliant work of Galton, strove to emulate his successes in another field. Having accepted the principle of individual differences, psychologists set about devising means of measuring them. They directed their efforts toward mental tests because the mind, to them, was the supreme organ for registering such differences. Now, they argued, if Tom differs from Dick and the difference is enregistered in the mind, let us measure the mind. But mind presented so many points of attack that the battle has been waged unceasingly to this day. We shall briefly review the mode of attack and the objectives aimed at.

The word "test" came into our educational vocabulary through the writings of Cattell, who in 1890 wrote on "Mental Tests and Measurements," arguing for the standardization of tests for individual differences. For Cattell mentality wore largely a sensory and sensory-motor aspect, so that a person's sensory and sensory-motor equipment became an index of intelligence. Bourdon (1895) viewed intelligence in terms of perception. The rate and time of perception became for him an index of mentality, and his cancellation test (canceling one specific letter on a printed page) found wide acceptance. So with Kilpatrick and Norsworthy. The former with ink spots and the latter with more elaborate material attempted to find a one-dimensional measure of intelligence in terms of some mental function.

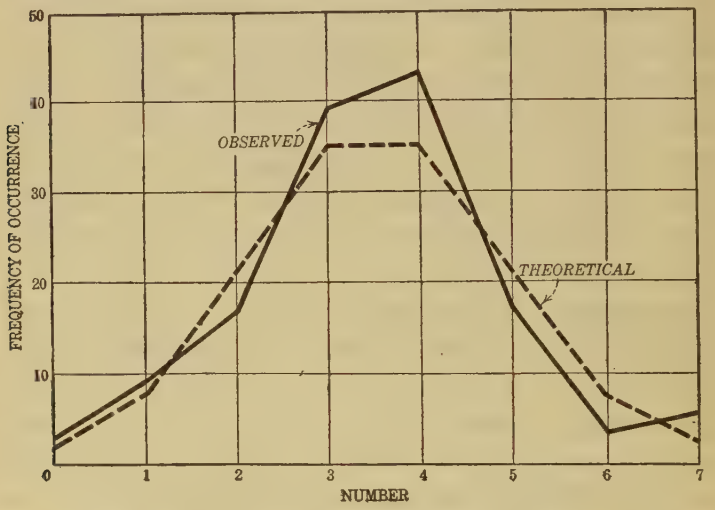
A more direct approach to the measurement of in-

telligence was made by Alfred Binet. But before this he had wandered far afield in search of tests that might measure a representative and all-inclusive function of intelligence. In 1896 he contributed an article in the *Revue Philosophique*, entitled "Measurement in Individual Psychology," in which he poses the question, "How can we measure richness of inspiration, accuracy of judgment, and the general ability of the mind?" He did not thereafter march straight to the goal of scale construction, but loitered on his way in the attractive side-paths of nutrition and anthropology. He thought he had found a one-directional road to intelligence by measuring "attention and adaptation."

In 1903 Binet struck the right road. In his book *L'Étude Experimentale de l'Intelligence* he equates intelligence with higher mental powers. The examinees were his two daughters. He tried such tests as the completion of sentences, the description of an object, and a description of a picture (a peddler and his boy pulling a wagon). He found other investigators laboring in the same field, collaborated with some, and criticized others in his *Année Psychologique*. His chief collaborators were Henri and Simon. Most of the contributions centered about the subject of abnormal psychology, in which branch of investigation the French have always excelled. Binet's interest in feeble-mindedness soon waxed warm, and with Simon he launched into a favorable critique upon the work of

Drs. Blin and Domaye in their attempt to devise objective tests of the subnormal mentality.

From an examination of the work of others to the construction of his scale was but a step. The so-called 1905 scale was soon published, followed by the 1908 and the 1911 scales. The subsequent scales represent



Probability Curve (Theoretical and Observed)

revisions in response to criticism, and embody a number of principles that have been accepted in scale making. The first principle is the Galtonian quantitative concept, that human mentality represents a continuous line of developments merging imperceptibly into each other — the idiot into the genius. It put to rout the dissidents who contended that idiocy, normality, and genius are distinct types of mentality, answerable to

characteristic impulses and reacting to the environment in unique ways. The second principle relates to the function of a scale. Binet maintained that mental measurement is not a mathematical process, that it is but a classification of individuals with reference to each other, a rank order of grading humanity. The third principle considers content. Binet strove to make his tests simple, heterogeneous, and non-pedagogical — a rich and varied sampling of tasks that do not draw too pointedly upon school learning. The fourth principle represents Binet's conception of intelligence as the unitary function of many psychic powers, capable of complete or partial concentration in any direction, possessing no one-dimensional common denominator and manifesting differences in ability when reacting on determinative levels. The fifth principle asserts the adequacy of mental age as a measure of intelligence. Binet, however, underestimated its predictive significance. He deprecated the use of scores, saying on one occasion that "this child who has passed through the rolling mill [test] comes before us with a certain total of marks, 36, for instance, or 70. We understand that 70 is nearer normal than 36, and that is all. We have no precise notion of the mental level of these candidates, no notion of what they can or cannot do. Did the one who obtained 36 have any comprehension of abstract ideas? We do not know and cannot divine."

Examination of the Binet tests need not detain us

here. We shall confine ourselves to a consideration of the subsequent "trends" in mental testing, noting particularly (1) the Binet revisions, (2) the statistical manifestation of mental scores, (3) the extension of Binet's technique to group tests, and (4) the achievement tests.

1. Binet Revisions. After the final revision of the scale (1911), criticisms poured in unceasingly. Vaney gave the tests to subnormals. He also used the scale for diagnosing feeble-mindedness. He found it a highly satisfactory instrument. Mills, Descoedres, D'Allones, Cruchet, and Lapie added their testimony. Some of it, however, was not favorable, and the opinion gained ground that the scale was too easy and that the tests were too few for children of younger years. The most damaging criticism came from two Milanese psychologists, Treves and Saffiotte, who had already begun an evaluation of the Binet Scale in 1909. In 1912 they published the result of their studies in the *Année Psychologique* at Binet's request. They had examined some 600 pupils from the six grades of the elementary schools of Milan. A preliminary study soon convinced them of the inadequacy of the scale. They therefore rearranged the tests, placing them in the order suggested by degree of difficulty experienced by the examinee. They found that the scale can best be used as a qualitative measure delineating these large groups: dull, average, and bright. They contended that, as race, age, and social status affected the final standing of the examinees, there

should be three scales, each standardized on racial, chronological, and social basis; that Binet's hypothesis of "pure intelligence" was untenable. They insisted on the reciprocal relation of mentality and environment. "What is this intelligence," they asked, "independent of culture, which the scale pretends to measure? Is it an intelligence which develops itself aside from all external influences? The intelligence, as the supreme expression of the total personality, demonstrates, in the last analysis, the capacity to adapt oneself to the social life."

Adaptations of the scale were soon made in many countries. In America the pioneer work was done by Goddard at the Training School for Feeble-minded Children at Vineland, N. J. Goddard visited France, learned something of Binet's work, and speedily became convinced of its usefulness. In the revision he used the 1908 version. From Vineland the scale spread to other institutions and was soon used in classifying normal children as well as feeble-minded. Because the scale was first used in testing the feeble-minded, it aroused in the laity a feeling that its employment among normal children would be socially prejudicial. This feeling has not entirely abated.

Other revisions soon followed, notably those of Terman, Kuhlmann, and Herring. The Terman revision, which appeared in 1916, is standard. It represents five years of exhaustive study. While it uses the Binet Scale as its groundwork, it may be considered a new

test. As Pintner remarks, " Much of what is only implicit in Binet is made explicit by Terman." And it marks one more advance. It applies Stern's conception of intelligence quotient (the ratio between chronological age and mental age measured in months), affording the tester a more dependable measure of mentality.

What characterizes other intelligence scales is the introduction of new tests and new methods of scoring. The Point Scale by Yerkes, Bridges, and Hardwick (1915) attempts to give greater accuracy to I.Q. by assigning part credits to answers instead of following the " all or none " method.

The Pintner-Paterson Performance Scale (1917) differs radically from the Binet Scale. The influence of linguistic proficiency is minimized, thus removing a handicap from the foreign born or those not linguistically endowed. The scale is used as a supplement to the Binet Scale in neutralizing the effect of verbal deficiency.

The Herring Scale (1921) is a Binet adaptation, using different methods of scoring and of computing mental age. The Kuhlmann Scale and the Burt Scale will be discussed in another chapter.

2. Statistical Elaboration of Scores. Binet's conception of mental age was purely qualitative. It was an arithmetical expression of a rank order. As Kimball Young phrases it, " While he [Binet] attempts standardization, at least in the 1908 and 1911 revisions, his idea seems ever to express by the ingenious notion of mental

age the real qualitative difference between pupils." " Too frequently," Young warns, " in our recent craze for the statistical refinement of his and other scales we have tended to forget his own valid criticism of just this danger." But the statistical current cannot entirely be stemmed. Indeed it is too valuable a method to be ignored. We shall see how profitable it may become when coupled with clear psychological thinking.

The impulse to mathematical evaluation of mental data came from Galton and was intensified through the efforts of Pearson and Spearman. To them is due the credit of formulating short methods of computing correlation. Correlation is a one-to-one relation between two scores designed to test the presence of a common factor operating to produce both. A pupil who receives a grade of 100 in spelling and 100 in composition is said to show a positive correlation of 1.00 between composition and spelling, owing to the possession of a common linguistic fund. On the other hand, a score of 100 in arithmetic and 0 in drawing shows a negative correlation, arguing for the wide disparity between the two abilities. Correlation may also obtain between the scores of a number of individuals, thus affording an opportunity for comparing the relative value of two different scales.

Carl Spearman precipitated a conflict in 1904 when he announced the statistical proof of the presence of two factors in intelligence. Having sampled the vari-

ous school abilities — arithmetical, linguistic, scientific, and manual — he calculated their interrelations and found that the correlation of each with the others formed a hierarchy. From this he argued that the positive correspondence pointed to a common factor operating in all mental processes and a specific factor functioning in each of the distinct mental abilities. A boy is clever in arithmetic because he has good general ability and a high order of specific mathematical ability; and so for other subjects.

The controversy is still raging and shows no signs of abatement. Webb, Garnett, and Burt may be said to side with Spearman, Burt inclining to the notion that this common factor is general intelligence; that it is innate; and that it is a unit character responsive to Mendelian laws. On the other side are ranged Thomson and Thorndike. Thomson repudiates the statistical validity of Spearman's proof, asserting that a hierarchy may be produced by chance, as in dice throwing, without the intermediation of a common factor. Thorndike does not take a purely statistical attitude. He dissents on the ground of low correlation between the level of sensitivity and the level of association, pointing to a weakening of the supposititious common factor. He concludes instead that the " mind must be regarded not as a functional unit, nor even as a collection of a few general faculties which work irrespective of particular material, but rather as a multitude of functions each of which involves content as well as form, and so is related

closely to only a few of its fellows, to the others with greater and greater degrees of remoteness."

The application of the statistical method, though of inconclusive force in the solution of the problem of the nature of intelligence, is of decided usefulness in rendering measuring more objective and more precise. Correlation is employed in proving the validity of new testing material by subjecting examinees at the same time to tests of known discrimination. The relative permanence of the I.Q. redounds to the benefit of mental testing. The use of conversion tables holds the promise of multiplying testing material of wider appeal owing to the ease of converting the scores made into standard Binet I.Q.'s. Again, the comparison of groups is facilitated by employing standard measures of scatter and variability, thus indicating the relative distribution of scores.

However, educational statistics do not always follow mathematical principles. Correlation is not an infallible way of proving the validity of a test. The fact that individuals examined by National Intelligence Tests test as high in the Binet–Stanford Scales permits the assumption of similarity, but not of identity. Nor can the I.Q. formula be extended to other relations arising by mere analogy to substantiate the argument that, since the I.Q. is acceptable, the E.Q. and A.Q. are just as valid. The E.Q. states a relation between educational age and chronological age, while the A.Q. expresses the still remoter relation between E.Q. and

I.Q. True, the formula $\frac{\text{E.Q.}}{\text{I.Q.}} = \text{A.Q.}$ wears a mathematical expression. But it has yet to be proved that the accomplishment quotient (A.Q.), compounded out of chronological age, mental age, and educational age, has in the nature of intelligence a predictive value. The various quotients are ratios, but they are essentially qualitative and not quantitative in their application.

3. Group Tests. The individual feature of the Binet Scale was bound to irk those not engaged in clinical diagnosis. Thorndike and others soon adapted mental testing to group requirements. Pyle gave tests to a group of children in 1913. Pintner followed in 1917. The greatest impetus to this tendency came from the World War (1917–18). It happened that Otis and Terman were hard at work on a group intelligence test (though Lobsein of Germany worked on group tests as early as 1914). This was accepted as the nucleus for a more comprehensive test, and with the aid of other psychologists the Army Tests came into being. They are known as the Army Alpha and the Army Beta. These tests served as models for numerous tests that are still making their appearance. The Army Alpha influenced the composition of all linguistic tests, borrowing considerably from other psychologists, notably Ebbinghaus, Zihen, and Binet. The Army Beta is a nonlinguistic test, depending for its technique on pantomime and demonstrations on the blackboard. The procedure is acted out before the examinees.

4. Achievement Tests. The interest in the objective measurement of educational subjects antedates in this country the introduction of the Binet tests. It quickened with the publication (1879) of Rice's report on spelling. The leadership soon passed over to Thorndike, who has maintained it unchallenged to this day. His *Introduction to the Theory of Mental and Social Measurement* (1904) directed attention to a virgin field for investigation. In quick succession appeared Stone's Arithmetic Tests and Thorndike's Handwriting Scale. The publication of educational tests and scales continued in full flood. Just as the introduction of mental tests stimulated interest in psychology, the devisal of educational tests has been conducive to a reorientation of methodology. The curriculum is being analyzed from the point of view of content and power. The educator is making a sharp line of demarcation between subject matter that is fundamental and content that is derivative therefrom; between bonds that must be cemented in a conscious act of learning and bonds that may grow through transfer of learning. In the construction of educational tests the psychologist displays great concern over the growth of power — the ability to apply fundamental processes and principles to new situations. Whence the equal emphasis on informational tests and reasoning tests.

That educational tests are growing in popularity is evidenced by the number of tests available in every subject. To enumerate a few:

Monroe Timed Sentence Spelling Tests
Charters Diagnostic Language Tests
Briggs English Form Test
Pressey Diagnostic Tests in English Composition
Hillegas Scale
Hudelson Scales
Hahn-Lackey Geography Scale
Cleveland-Survey Arithmetic Tests
Monroe Diagnostic Tests in Arithmetic
Monroe Standardized Silent Reading Tests
Thorndike-McCall Reading Scale
Holtz-Godsey Latin Teaching Tests
Handshin Silent Reading Tests for French
Minnick Geometry Tests
Douglass Standard Diagnostic Tests for First-Year Algebra

There are still other tests that partake of the nature of intelligence tests and achievement tests. They may be termed mental-educational tests because they combine, in one booklet and in a single ratio, the mental and educational indices of pedagogical efficiency. Thus the Pintner Mental-Educational Tests test intelligence adequately. They also cover the principal subjects, reading, arithmetic, grammar, history, and geography. The author furnishes standard tables for the purpose of converting scores into a mental and an educational index, data which inform the teacher whether any pupil is working to capacity. Other tests will be discussed elsewhere.

(4) *Biological Qualitative Phase*

Seguin's warning, that man, being a unit, functions as unit, did not fall on deaf ears. Binet's conception of intelligence was also dominated by the biological viewpoint that the individual reacts with the total personality. Moreover, he eschewed the purely quantitative expression of intelligence, proposing instead a qualitative classification in the form of an arithmetical expression (mental age). Writing of the Binet Scale, Binet and Simon declared that " this scale, properly speaking, does not permit the measure of intelligence, because the intellectual qualities are not superposable and therefore cannot be measured as linear surfaces are measured, but are on the contrary a classification, an hierarchy among diverse intelligences; and for the necessities of practice this classification is equivalent to a measure."

Let us consider for a moment the difficulties which beset the psychological quantitative evaluation of individuals. First of all, mind is an integrative unity of psychic powers functioning in varying combinations. Observing, imagining, judging, and reasoning, while not distinct entities, are sufficiently distinctive to merit differentiation. Yet their countless permutations resist the application of infinitesimal calculus: it is all but impossible to calculate the contribution of each one to any psychological act from spelling a word to solving an arithmetical problem. On the

other hand, the assumption that reasoning hierarchizes all the mental powers offers no escape from the dilemma. You cannot reason without observing, remembering, imagining, and judging, but a defect in reasoning may be traceable to a defect in any one of these. For the purpose of evaluation, logical proficiency is as conclusive in determining mental status as the condition of bodily vigor is in diagnosing the state of an individual's health.

We err no less in tacitly considering intelligence the chief index of personality. Indeed, the tendency of statistizing intelligence leads to its overvaluation; so the I.Q. is deemed in many quarters highly predictive of future promise or lack of promise. This position may be maintained if we comprehend under the term " intelligence " other personality factors. That this is not the case is shown by the nature of the testing material entering into the make-up of most scales.

We may divide this material into seven classes:

1. Questions calculated to establish the degree of orientation attained. " What is your name? " " Which is your left eye? "
2. Tasks designed to test manual dexterity. Copying a drawing, judging weights
3. Tests ascertaining the content of the mind. Information questions of all kinds
4. Tests requiring a degree of adaptation. Problems involving new relations
5. Tests revealing the " ability to take and maintain

a mental set." This includes commands to be executed, etc.

6. Tests disclosing "the power of auto-criticism." Resistance to suggestion, criticism of absurd statements
7. Tasks indicative of the quality of memory. Memorizing sets of numbers and sentences

Classes 4, 5, and 6 are the only ones that plumb other than intellectual depths. Their number is relatively small; and while they do test perseverance and emotional control, they do so incidently and, like the other types of tests, evaluate a reactive phase.

The normal condition of the organism is one of flux; the mind is in constant movement between the unconscious urges and environmental trends. At any instant of inward or outward stimulation a condition of momentary concentration is brought about involving the personality forces released into the activity stream. In time this mercurial personality comes to occupy, in its fundamental reactions, more stable positions, and the self functions under settled character traits. We may call this self the spontaneous personality — the personality set in motion under normal conditions of stimulation. But this fundamental personality is undergoing ceaseless change occasioned by the ever-changing environment, and the individual becomes the theater of a struggle between the forces of habit and those of progress.

Now it is conceivable that the mental test may find

the examinee in the throes of a revolutionizing adjustment in which the predominance of an instinctive or emotional mood may obscure the real self. Under such circumstances the intelligence test may yield an index of reactivity — the mental make-up as it then functions — and not an index of spontaneity — the habitual self. True, psychologists attempt to obviate the difficulty by giving alternative tests, one for each of two successive days. But even then the sampling may bring to light the slightly modified reactive self.

Neither the psychological nor the quantitative approach is therefore adequate. Psychology has tended to give personality a decidedly mentalistic aspect, and statistics are verging toward a linear conception of mentality. Young's criticism, though severe, is justified. "We are rapidly losing our older notions of quality, of calm and divergently integrated personalities. If the older laboratory psychology produced a psychologic man who had no individuality, the present statistical treatment of him in terms of I.Q. and scores, buttressed by averages of correlation quotients, tends to make man a psychological robot with no emotion, no quality, no personality."

The inadequacy of the I.Q. is becoming apparent. Educators have been hard put to explain the vexing discrepancy between an individual's I.Q. and his performance. Teachers are demanding a system of evaluation that will appraise the total personality, so that they may be spared the disappointments

arising from attempting to bridge the gulf between high mentality and poor results. They insist on the creation of supplementary measures designed to plot all personality factors functioning in learning. The biological qualitative evaluation of personality is the answer.

First of all, the biologic interpretation of personality views the individual in his totality and regards intelligence neither as supreme expression of personality nor as the prime mover in all learning. Rather does it rank intelligence as the mediator between instincts and motives and the environment. It considers that the unconscious impulses, having become crystallized into desires and wishes, still retain their unconscious nature, but operate none the less as mainsprings of conduct; that intelligence which manifests itself in school life is but a phase of this governing intelligence relating itself intimately to the unconscious desires; and that in any act of learning there enters not only intellection, but instincts and emotions which ever after form its affective tone.

The biological phase of the measuring movement diverges from the parent movement in the content of the tests and in the abandonment of too fine a quantitative manipulation of scores. That is not a new movement. Clinical psychologists in all countries, in France in particular, have always emphasized the " total adjustment of the organism personality." The Germans also devised scales for measuring it. In 1915

Rossolimo attempted a clinical picture of the individual. He called it a "psychograph." He used the older categories of will, attention, perception, memory, comprehension, etc., but believed in a unified personality. Each personality trait was judged on the basis of ten, and judgment was arrived at through tests and the teacher's judgment. The marks were then plotted, and the curve constituted the individual's psychograph. These psychographs were later embodied in the form of observational records, which under comprehensive headings furnished information about the individual, who remained under observation for long periods.

In this country the trend has been in the direction of psychographing in units of personality. With tests of emotional stability, of character, of perseveration, Doll has given a picture of the personality in the form of an exhaustive "case" history. So has Mateer. Bronner (1914) has constructed a test designed to disclose moral delinquency, while Kohs (1922) has devised an ethical discrimination test. Green and Jung urge the unsymbolic interpretation of day dreams, asserting that the dream life of a child holds the key to his secret desires and longings. Watt and Terman use a modified form of Jung's word-association test to uncover insidious complexes and conflicts.

Finally, mention must be made of two distinctive tendencies, a behavioristic evaluation of personality and a biological quantitative interpretation of personality traits.

John Watson speaks for the Behaviorists. He would demand a cross-section of the total personality organization of the individual, which he envisages as a system of built-in habits. He would seek information, in qualitative form, of the nature of the individual's vocational habits, mental habits (including the I.Q.), also systems of habits (commonly known as sentiments and attitudes), such as religious habit system, emotional habit system, personal habit system, and recreational habit system. In other words, he insists upon a complete report of the life history of the individual.

The other extreme is represented by Barbara S. Burks's Scale of Promise (1925). Miss Burks derived her inspiration from Terman's pioneer study of over 1000 gifted children (1921–25). Terman's statistical conclusions, however, have been based on 643 children, who will be followed up for a number of years. The data have been gathered from many sources, so that the total personality has been surveyed. Terman's survey is illuminating. For each child the following sources were tapped:

1. Two intelligence tests (Stanford-Binet and National)
2. A two-hour educational test (Standard Achievement)
3. A fifty-minute test of knowledge and of interest in plays, games, and amusements
4. A fifty-minute test of general information in science, history, literature, and art

5. A battery of character tests, including two overstatement tests, a test of reliability of performance in face of temptation not to follow rules laid down, a social attitudes test, and a test of emotional stability
6. A four-page interest blank filled out by children
7. A word-association test of intellectual, social, and activity interests
8. A two-month reading record kept by the children
9. A complete medical examination
10. A sixteen-page Home Information Blank filled out by the parents, including ratings on twenty-five traits
11. An eight-page School Information Blank filled out by the teachers, including ratings on the same twenty-five traits
12. When possible, ratings of the home on the Whittier Scale of Home Grading

While this examination is exhaustive, it is not pretentious. Hollingworth is not overcritical when she states that " the complexity of a human mind is such that it can never be assayed by the application of a single test." Miss Burks, however, was interested in a Scale of Promise — a predictive scale of future success. She, therefore, chose her tests with a view to their prognosticative value and not primarily to their diagnostic discernment. Relying chiefly on the judgment of two trained psychologists, who assigned to

each of the items a rating commensurate with their supposed influence on a successful career — successful in an individual as well as a social sense — Miss Burks chose ten items * of predictive import:

1. Persistence and mental drive
2. Moral outlook and standards
3. Interest in a vocation in a special field
4. Social adaptability
5. Intelligence quotient
6. Health and physical energy
7. Social status of family
8. Interest in activities
9. Ability to profit from school work
10. Originality

While the outlook is decidedly biological — adequacy and richness of adaptation — the system of marking is quantitative. Herein lies the chief weakness of the scale. The subjective factor has not been eliminated; the resort to an arbitrary rating of 7 for each of the ten items, the weighted value of each, and the correction by dividing by the standard deviation (S.D.) import into the scale too many approximative quantities. "A child's final score on any item was equal to his rating (on a scale of 7) times 'importance' value divided by the S.D. His final promise score was equal to the weighted total of all the separate (10) scores."

* "A Scale of Promise," *Pedagogical Seminary*, September, 1925.

Summary

The testing movement began as a qualitative evaluation of the degrees of defect, applied for purposes of social segregation. The growing interest in individual differences as determined by heredity and manifold powers of mentality, coupled with the vogue of statistical manipulation of social data, soon led to a quantitative appraisal of what was comprehended under the term "intelligence." A reorientation to the meaning and significance of intelligence came as a revolt against a too rigidly quantitative interpretation of a mentalistic personality. Intelligence was resurveyed and assigned to a coördinate rôle with other personality traits, notably emotional stability and persistence. Its complexity militated against its being compressed into a mathematical formula. The biologic trend set in. There is a marked tendency to-day to plumb the total personality; and while statistical treatment is desirable, trained judgment directed to specified items of personality value are deemed superior to a too mechanistic objectivity.

Exercises for Self-Testing and Investigation *

I. With respect to the following psychologists one of the three phrases refers to their most important contribution to the

* These exercises and those at the ends of following chapters are planned to test the student's knowledge of the content of the chapter they follow and to serve as a guide for investigation of allied fields of education. Such investigation may be conducted by means of the texts

measuring movement. Indicate whether it is the one marked (*a*), (*b*), or (*c*).

1. Binet:
   (*a*) Intelligence scale
   (*b*) Father of Intelligence Testing
   (*c*) Mental age
2. Terman:
   (*a*) Author of *Measurement of Intelligence*
   (*b*) Author of *Genetic Study of Genius*
   (*c*) Revision of Binet Scale
3. Thorndike:
   (*a*) S–R formula
   (*b*) Studies in animal psychology
   (*c*) Use of statistics in testing
4. Seguin:
   (*a*) Author of *Idiocy: Its Diagnosis and Treatment*
   (*b*) Statement that causal factors of idiocy are psychological
   (*c*) Unitary view of individuality
5. Galton:
   (*a*) Introduction of scientific method into the field of associative learning
   (*b*) Use of mathematical methods of Gauss and Quetelet
   (*c*) Study of heredity

and "studies" named in the Bibliography and by making original studies of tests and scales. Most tests and scales may be obtained by writing any or all of the following publishers:

Bureau of Educational Research, University of Illinois, Urbana, Ill.
A. S. Courtis, 1807 East Grand Boulevard, Detroit, Mich.
J. B. Lippincott Co., Philadelphia, Pa.
The Public School Publishing Co., Bloomington, Ill.,
Teachers College, Columbia University, Bureau of Publications, N. Y. C.
The World Book Co., Yonkers, N. Y.

6. Stern:
   (*a*) Use of correlation in tests
   (*b*) Concept of I.Q.
   (*c*) Definition of "intelligence"

7. Cattell:
   (*a*) Argument for the standardization of tests of individual difference
   (*b*) Use of sensory-motor data as an index of intelligence
   (*c*) Introduction of term "test" into psychological measurements

8. Goddard:
   (*a*) Adaptation of the Binet Scale
   (*b*) Studies in feeble-mindedness
   (*c*) Studies in results of the Army Tests

9. Spearman:
   (*a*) Statistical formula of correlation
   (*b*) Intercorrelation formula
   (*c*) Theory that two factors enter into intelligence

10. Franzen:
    (*a*) Deviser of scales
    (*b*) A.Q. formula
    (*c*) Formulation of a theory of personality

II. In the following exercise appear a list of authors and a list of their writings. The student is required to match authors and books.

AUTHORS

1. Seguin
2. Thorndike
3. Galton
4. Cattell
5. Terman
6. Watson
7. Fitzherbert
8. Binet
9. Mateer
10. Stern

BOOKS

1. *Mental Tests and Measurements*
2. *English Men of Science*
3. *Psychological Methods of Testing Intelligence*
4. *The Intelligence of School Children*
5. *New Natura Brevium*
6. *Année Psychologique*
7. *Idiocy: Its Diagnosis and Treatment*
8. *Behaviorism*
9. *The Psychology of Algebra*
10. *The Unstable Child*

III. Tests and scales have been classified into:
   A. Social qualitative
   B. Psychological qualitative
   C. Psychological quantitative
   D. Biological qualitative

Arrange the following measuring instruments in accordance with the above classifications; indicate right choice by affixing A, B, C, or D after each type.

1. English norms for measuring sanity
2. The National Intelligence Tests
3. Observational records
4. Psychiatric tests
5. Psychographs
6. Jung's classification according to types
7. The questionnaire
8. The Binet Scale
9. Free association tests
10. Legal definitions of insanity

BIBLIOGRAPHY

Burks, B. S. — "A Scale of Promise," *Pedagogical Seminary*, September, 1925

Burt, C. — "Experimental Tests of General Intelligence," *British Journal of Psychology*, Vol. 3, pp. 94–177, 1909–1910

Cattell, J. McK. — "Mental Tests and Measurements," *Mind*, Vol. 15, pp. 373–380, 1890

Fernald, W. E. — *The History and Treatment of the Feeble-minded*, Twentieth National Conference of Charities and Correction, 1893

Galton, F. — *English Men of Science*, Appleton, 1890

—— *Hereditary Genius*, Appleton, 1880

Hart, B., and Spearman, C. — "General Ability, Its Existence and Nature," *British Journal of Psychology*, Vol. 5, pp. 51–84, 1912–1913

Herring, J. P. — *Herring Revision of the Binet-Simon Tests*, World Book, 1922

Johnson, G. E. — "Contribution to the Psychology and Pedagogy of Feeble-minded Children," *Pedagogical Seminary*, Vol. 3, No. 2, pp. 246–391, 1895

Kilpatrick, E. A. — "Individual Tests of School Children," *Psychological Review*, Vol. 3, pp. 274–280, 1909

Norsworthy, N. — "The Psychology of Mentally Deficient Children," *Archives of Psychology*, No. 1, 1906

Peterson, J. — *Early Conceptions and Tests of Intelligence*, World Book, 1925

Pintner, R. — *Intelligence Testing*, Holt, 1923

Pressey, S. L. and L. C., *Mental Abnormality and Deficiency*, Macmillan, 1926

Seguin, E. — *Idiocy: Its Diagnosis and Treatment by the Physiological Method*, Teachers College, Columbia University, 1907

Sharp, S. E. — "Individual Psychology," *American Journal of Psychology*, Vol. 10, pp. 329–391

Stern, L. W. — *Psychological Methods of Testing Intelligence*, Warwick and York, 1914

Symposium — "Intelligence and Its Measurement," *Journal of Educational Psychology*, Vol. 12, Nos. 3 and 4, 1921

Terman, L. M. — *The Intelligence of School Children*, Houghton Mifflin, 1919

—— *The Measurement of Intelligence*, Houghton Mifflin, 1916

Thomson, G. H. "General versus Group Factors in Mental Activities," *Psychological Review*, Vol. 27, No. 3, pp. 173–190, 1920

Thorndike, E. L. — *An Introduction to the Theory of Mental and Social Measurements*, Columbia University, 1904

Watts, F. — *Abnormal Psychology and Education*, Appleton, 1924

Webb, E. — "Character and Intelligence," *British Journal of Psychology*, Monograph No. 3, 1915

Whately, S. — *The Measurement of Emotion*, Harcourt Brace, 1922

Witmer, L. — "Clinical Psychology," *Psychological Clinic*, Vol. 1, Nos. 1–9, 1907

Young, K. — "The History of Mental Testing," *Pedagogical Seminary*, Vol. 31, No. 4, pp. 1–48, 1924

CHAPTER II

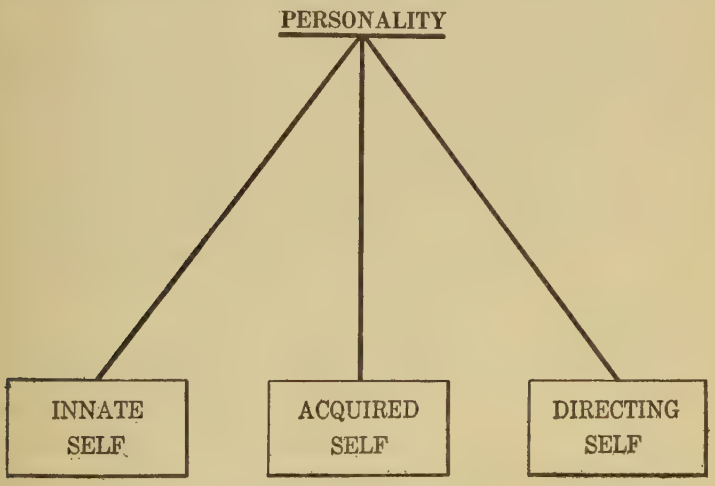
AN ANALYSIS OF PERSONALITY

The human personality is a highly complicated machine, and a definition of personality must include many terms. The multiplicity of such terms is due to the many-sidedness of personality — the social, the psychological, and the psychiatric. Society insists that its members must be socially effective, that they must learn to work coöperatively and individually. Any deviation from socially valid standards, whether in the field of concerted action or individual endeavor, is visited with displeasure or punishment.

Social standards vary from age to age. There is, however, general agreement in stressing the attainment of balance between biologic and environmental forces. Man is a living organism with ready-made structures and innate life impulses; he is molded by environment into a social being. But the molding process must proceed along the lines indicated by nature; you can shape forces only when you follow the direction of their movements. The moment you go counter to their dominant tendencies you are bound to meet resistance. You cannot pour the individual's life forces into a mold

as one pours concrete into a form. For one thing, human nature is too fluid; educability bases itself upon this quality. For another, the environment is not static for any individual; it will not harden into "forms."

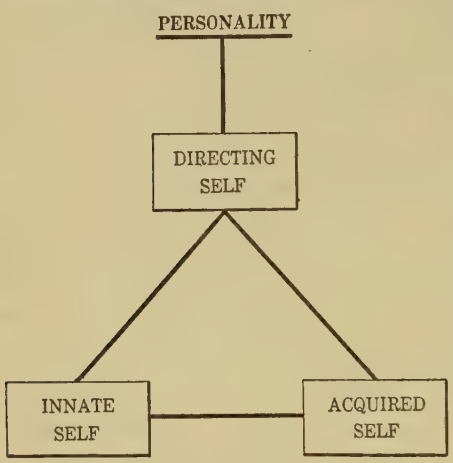
How is personality fashioned? In childhood and adulthood in much the same way as a photographer



makes a picture. The camera and the sensitized plate correspond to the innate equipment of the individual. The sun, clouds, and shadows constitute the environment. But the "hypo" bath is administered by the environment (the photographer) in infancy and by the acquired equipment in childhood and adulthood. At birth and for some time thereafter, the infant is assailed by stimuli, which, impinging upon the sensitized

instinctive " plates," rapidly become " developed " through their own activity. Later, intelligence organizes impulses and trends into a " self " which directs this " developing " process.

Personality may now be envisaged as a resultant of three factors: the innate self, the acquired self, and the directing self. This threefold division is largely schematic and will be modified as the discussion proceeds. It is, however, comprehensive enough to serve as a frame. The interaction of the three elements produces the product known as personality. It may be defined tentatively as the environmental and internal integration of innate and acquired forces into a complexly organized mechanism of sufficient flexibility to enable the individual to adjust himself to his entire environment or to any segment of it.



A. Innate Equipment

The innate equipment of the individual is physical and psychological.

(1) *Physical Innate Equipment*

1. Cellular Structure of the Body. The physical bases of life are built upon cells. A cell consists of a body with many appendages. It has a nucleus. The body gives off branches: one kind, tree-shaped, called dendrites, and another branch, straighter and more slender, called the axis-cylinder. The complete cell is known as a neurone. The cells are interconnected, the axis-cylinder of one leading to the dendrites of another. The neural impulse is conveyed through dendrites and cell body to other cells by way of the axis-cylinder. The circuit is then: dendrite ⟶ cell body ⟶ axis-cylinder ⟶ dendrite. Conduction is always one way and always forward.

There are four different kinds of cells, each with its characteristic protoplasm, nucleus, and chromosome. The whole is held together by a membrane. These cells form tissues which enter into the formation of every organ of the body.

(*a*) *Tissue for Covering the Body and for Lining All Openings.* Cells of this character form the epithelial tissue from which, by modification, nature fashions the layers of the skin, the teeth, the eyeball (cornea), the tear glands, the sweat glands, the salivary glands, and many other glands. It also enters into the formation of the sensitive parts of the sense organs. In addition, it serves as a lining for inside tubes and cavities — the alimentary tract, the blood vessels, and the openings in the brain (spinal canal).

(*b*) *Tissue for Connecting Parts of the Body.* This material is known as connective tissue. It builds a framework for the bones (periosteum), tendons for the attachment of muscles, cartilage for the nose, and strong fibers (white fiber cartilage) for holding together movable bones. There are four varieties of connective tissue, cartilage, bone, elastic tissue, and fibrous tissue.

(*c*) *Muscular Tissue.* The body must move about. It needs a substance that will, by contraction and relaxation, produce locomotion. There may also be motion without direction. Nature has therefore provided muscular tissue and nervous tissue. Of the former, the striped variety enters into the structure of the muscles and the tongue. It is used whenever the body needs a mechanism for moving large masses or for moving with great rapidity. The unstriped or smooth tissue forms layers which coat the stomach, the intestines, the bladder, the iris of the eye, the arteries, and the veins.

(*d*) *Nervous Tissue.* Psychologists have placed great emphasis on the nervous tissue because of its dynamic properties. In a way it forms the cement that brings the bodily structures into coöperation. But it is becoming increasingly accepted that in the process of adjustment to the environment the other tissues act originally and modifyingly upon the stimulus. Thus, by analogy, if the stimulus is the button, the nerve the electric wire, and the muscle the bell, pushing the button rings the bell; that is, there is muscular re-

sponse. But tension of muscle, the stress of muscle upon its supporting bone, may also initiate innervation (kinæsthetic sensation). The same is true of ductless glands, as we shall see later. In other words, the bell itself may, by mechanical manipulation such as bodily pressure, begin to ring. Nevertheless, the nervous structure is tremendously important, and will repay an extensive elaboration of its constituent elements.

2. Principal Organs of the Body. With the four types of tissue nature builds organs. The principal ones are (*a*) sense organs, (*b*) reacting organs, and (*c*) directing and coördinating organs.

(*a*) *Sense Organs.* These structures are known to the average reader. They may be summarized as follows:

Sense	Organ	Stimulus
Visual	Eye	Ether vibrations
Auditory	Ear	Air waves
Gustatory	Tongue	Sapid fluids
Olfactory	Nose	Gaseous particles
Dermal	Skin	Warm or cold objects
Temperature		Warm or cold objects
Pressure		Cutting, burning
Pain		Pricking
Kinæsthetic	Muscles, tendons	Stress in musculature

When a stimulus strikes an appropriate sense organ, there results some kind of physical and chemical change. This change takes place in the epithelial cells and in turn initiates a neural impulse (of whose

exact nature we are ignorant) in the nearest nerve ending. The impulse travels from neurone to neurone by way of the central nervous system and terminates in some muscle or gland.

(*b*) *Organs of Response.* Let us see now what takes place in the muscles. The muscles are so arranged that the tension of one takes place simultaneously with the relaxation of an adjacent one. Indeed, the muscles of the entire body are arranged on this give-and-take principle, and the simplest act, like picking up an object, involves a rearrangement of the entire musculature. When you button your shoe, you do it with your whole body.

The internal organs are formed of unstriped muscle. They are called collectively visceral organs. The viscera also are capable of responding. But what is more important for personality, they are also capable of initiating impulses: they are organs of stimulation. Consider, for instance, such organs as the stomach, lungs, heart, and intestines. They are incessantly active in an accelerated or retarded tempo. Their contents are constantly in motion. The stomach contracts rhythmically when it empties its contents. Hunger intervenes. This produces a restlessness culminating in search for food. This restlessness may set up other reactions, which may interfere with the individual's thinking and throw him into a state of disequilibrium. Many cases of class indiscipline and inattentiveness arise from such causes.

The endocrine glands. Of greater moment to the organism are the ductless, or endocrine, glands. We know that they affect the personality in many ways. The most important of these glands are (1) the thyroid and parathyroid glands, (2) the adrenal bodies, (3) the pituitary body, and (4) the pineal body.

The thyroid glands are located along the sides of the windpipe and just below the Adam's apple. They have two lobes connected by a bridge. Blood vessels supply them with blood, and nerve fibers connect them with the nervous system. They, therefore, have means of switching to the directing mechanism. Hence their inclusion among the responding organs. They secrete a substance called thyroxin, which contains 60% iodine.

In the normal person, the secretions are absorbed and mental equilibrium is sustained. When the glands function poorly or intermittently, there ensue marked differences in behavior. Growth is arrested (there is incomplete ossification of the bones); the reproductive organs do not develop; the skin thickens; and the hair becomes colorless. Pedagogical retardation sets in. When the thyroids are removed, the individual reacts similarly. Restoration of normal functions, however, follows the use of the dried thyroid gland of a sheep or small amounts of thyroxin.

Sometimes the thyroids become overactive. The whole bodily mechanism speeds up: blood pressure becomes high, irritability characterizes the emotional

set, and increased activity of the brain produces insomnia. The glands produce a result like that which occurs when a motor is " raced " continuously. When they undersecrete, their action is on a par with that of the motor which is badly carbonized.

In close association with the thyroids are the parathyroid glands. These four masses, about the size of a pea, are more closely bound up with the nervous system. Their removal brings death, preceded by alarming symptoms: muscular tremor, spasms, high temperature, rapid breathing, vomiting, and diarrhœa. It is believed that their secretion regulates the discharge of nerve cells. Thus they constitute another link in the neural chain.

The two adrenal bodies, or glands, are situated near the kidneys. They discharge into the blood stream a substance known as epinephrin or adrenin, whose presence has the effect of heightening the emotional tone in fear, rage, and pain. Their action is chemical and physical. They split up the glycogen which is stored up in the liver and change it into blood sugar. The blood carries the blood sugar to the muscles, which, refreshed, are prepared to meet the muscular demands of the body. The free adrenin in the blood performs another useful task: it causes the arteries to expand, thus permitting a free and more rapid flow of blood to the laboring muscles. The increased volume of blood washes away the waste products which muscular effort produces. The effect upon the body of this accession

of energy has tremendous survival significance. It brings up timely reënforcements when the need for them is extreme. The adrenin enables man to fight harder and longer, to run faster and farther, and to concentrate more intensely and absorbedly. On the other hand, this too ready access to an armory of weapons makes of man a direct actionist, with overpowering proclivities to rattle the sword of emotion at the slightest signs of environmental conflict: the admonitory voice of intellect receives scant courtesy in the counsels of feeling.

The pituitary body is located between the roof of the mouth and the brain. It consists of two lobes, the anterior and the posterior, each functioning in a characteristic manner. In disease the anterior lobe may oversecrete or undersecrete. When the secretions grow too copious, the bodily functions speed up and forced growth takes place, eventuating in gigantism. When it undersecretes or when it is entirely removed, emaciation sets in, accompanied by lowering of temperature, unsteadiness in gait, and diarrhœa. Death ensues in a short time. We do not know what chemical substance the anterior gland secretes. Evidence is accumulating that it controls the growth of the skeleton and the connective tissue.

The posterior lobe may be removed without killing the animal, but its loss or diseased condition brings about adverse changes in the way our food is energized (metabolism). The body becomes tolerant of sugar

and grows fat. The posterior lobe has a stimulating effect on all striped muscle, the kidneys, and the mammary glands. It coöperates with adrenin in the task of activating the muscles.

The pineal, the thymus, and the puberty glands control and regulate the sex life of the individual. The rejuvenating operations of Voronoff and Steinach have for their purpose the stimulation and restoration of sex activity. Since this book concerns itself primarily with the school child, the discussion of glands of postpubertal significance need not detain us here.

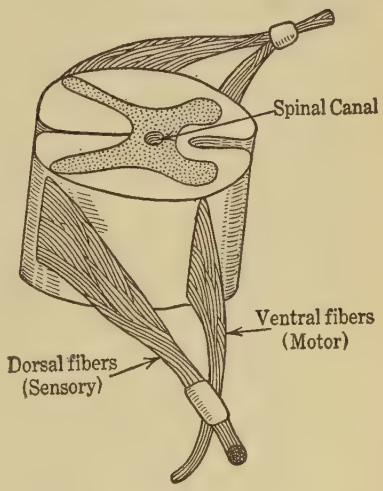
The workings of the entire glandular system are still at the stage of fragmentary discovery. Their introduction into the field of psychology is of very recent occurrence. Yet even medical men concede their importance in personality building. It is not an exaggeration to state that they interpenetrate the behavior of man from the simplest act of perception to the most complicated process of abstract thinking. We shall see that any act of cognition may produce oversecretion or undersecretion of the endocrine glands; that this hyperactivity and hyposecretion are as basic in the learning process as memory and imagination.

(*c*) *Conducting Organs.* The road from sense organ to reacting organ is paved with neurones. We shall find short cuts, detours, and bypaths. There are two systems of roads, the cerebro-spinal (main roads and detours) and the sympathetic (bypaths). Both systems are closely linked, so that a neural impulse can reach any

one of the reacting organs either by direct path or by devious detours.

The cerebro-spinal system is made up of the brain, spinal cord, nerve dendrites, and oxones, which lead to sense organs and to the sympathetic system. The brain, notably the cerebrum, is the highest organization organ.

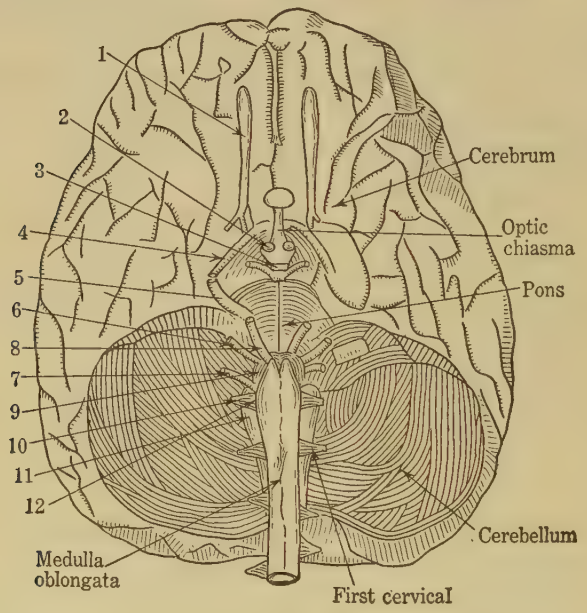
The sympathetic system consists of groups of nerves called ganglia. They are scattered throughout the body and are connected with the cerebro-spinal system. They function in situations in which contraction and relaxation of the unstriped muscle tissue (the heart, the alimentary canal, etc.) are called into play. They also regulate the secretions of the endocrine glands, which play so prominent but mysterious a part in emotional reactions and developmental processes.



Cross Section of the Spinal Canal

The shortest pathway from receiving organ to reacting organ is a short reflex arc. When you burn your hand, the reaction is immediate. The mechanism involved in this act consists (to simplify matters) of three neurones: an afferent neurone from the skin to the

spinal cord, a central neurone inside the cord, and an efferent, or motor, neurone leading to the muscle of the hand. The body contains longer pathways, extended for the purpose of delaying the reaction so as to permit deliberation, evaluation, and choice. Instead of pass-



View of Under Surface of Brain, with Nerves Numbered

ing through the spinal cord — the short unmediated cut to behavior responsive to the pressing need for immediate adjustment — the circuit turns sharply away to the brain, from which it issues, belatedly but decisively, to the efferent neurone. To be sure, the impulse travels a longer distance, but like a driver

allured by the inviting country road but not sure of his direction, it faces the mental traffic cop, submits to his mandates, and goes on its way with redoubled assurance of reaching its destination.

(*d*) *Other Bodily Structures Innately Adjusted to Environmental Forces.* It is freely admitted that human beings differ markedly in the structure of the various parts of the body, and that this variation produces measurable differences in personality. Environment but accentuates these differences. Training cannot make a singer of a man who has no adequate vocal equipment. Examples may be multiplied indefinitely. Some people begin life with a varied heritage of slender fingers, delicate throat structure, sensitive sense organs, and the muscular and skeletal equipment of prize fighters. Just as no two human beings have the same finger print, so no two persons have quite the same leg bones, the same glands. They differ also in brain and cord structure, in the construction of the circulatory mechanism, and in the length, breadth, thickness, and flexibility of the striped muscles. Each of these differences influences the course of personality development. Thus heredity injects another element to complicate the complex process of building personality.

(2) *Psychological Innate Equipment*

1. Instincts and Emotions. The cardinal principle of life is activity. This human organism we have

been describing is in a state of ceaseless activity. When conscious activity intervenes we have waking life; when conscious life is in abeyance we have sleep. Still activity persists. Some of this activity is innate.

(*a*) *The Instincts.* An instinct is a predetermined manner in which the organism acts. It has been defined as " a tendency to act in such a way as to bring about certain ends without having foresight of those ends." Note that it is an unlearned response, that it seeks adjustment, and that it is blind. William James enumerates climbing, imitation, hunting, fear, acquisitiveness, constructiveness, curiosity, sociability, shyness, love, jealousy, and many others. Recent writers, McDougall, Woodworth, Rivers, and Sherrington, have reduced these to four categories called " drives " — impulses, energy, sources, or motives. Their dissatisfaction with James's classification arises from the knowledge, largely experimental in character, that learned responses are initiated by the environment in early infancy, and that this increment of equipment soon becomes indissolubly bound up with instinctive behavior. The Behaviorists, notably John Watson, have clarified the situation still further and have shown, in a purely objective way, that the unlearned responses of infancy are few and definite, and that these are modified complexly by learned behavior to the obscuration of the original responses. Watson has found that the repertory of unlearned responses includes sneezing, hiccoughing, crying, early eye movements, smiling, manual

movements (hand, leg, trunk), feeding responses, crawling, the "extensor thrust" used in standing and walking, grasping, and blinking.

(*b*) *The Emotions.* Emotions are unlearned visceral responses to situations tending to endanger the individual's security or to add perceptibly to his well-being. They are as innate as instincts and share with them the responsibility of adapting the organism to its world. They differ in intensity of expression, in the nature of accompanying bodily changes, and in destination. An emotion calls out a rapid mobilization of those energies calculated to remove a menace; concentration is rapid, compact, and thorough. Bodily changes set off the signal: tremor of muscles, rapid respiration, pounding of heart, creeping sensation of the skin, and changes in the rate of glandular secretion. Instinctive behavior ceases when the organism is once more in harmony with the environment. The hand, for instance, is moved out of the way of the irritating pin point and the circuit is closed. In emotion, harmony is sought in the function of the internal organs, not in the external sources of stimulation. Instinct and emotion have a close bond in that the feeling of satisfaction or dissatisfaction follows instinctive activity. According to McDougall, emotion parallels instinctive trends: the instinct of flight and the emotion of fear; the instinct of pugnacity and the emotion of anger.

The mechanism of emotional activity has been

described succinctly by James in terms of the James-Lange theory. Briefly, we are afraid because we run; we do not run because we are afraid. The stimulus, a terrifying object, throws the bodily organs into a state of turmoil. Our perception of this condition constitutes the emotion. Psychologists are inclined to accept this theory, particularly its insistence on the preponderant influence of vaso-visceral-motor changes involved. Indeed, the emphasis currently placed on this constituent of emotional behavior is most pronounced. They do not follow James in his classification of emotions, but contend that this heart-gland-muscle reaction may be called out by a host of stimuli by a process of transference — a process which exposes the major organs of the body to the profound stimulation of the entire environment.

Psychologists will always disagree on the classification of emotions because their manifestations are endlessly varied and their component elements multitudinous. An attitude of arbitrariness is inescapable in dealing with fluid and evanescent data. Watson's * classification, however, has the merit of objective derivation. His method being genetic (developmental), he has a definite point of departure, deriving a comfort from the relative simplicity of infant behavior. He finds that the original emotional organization of infants is not intensive, and that its mode of reaction follows

*Watson, John B. — *Behaviorism*, W. W. Norton and Company, 1925.

a fixed stimulus — response formula. The primary emotions are three in number: fear, rage, and love.

Fear. This form of reaction may be illustrated as follows:

Stimulus	Response
Loud noises: striking of steel bar with a hammer; loud, shrill, hissing sounds	Respiratory pause, vasomotor changes, closure of eye, clutching of hands, puckering of lips complicated later by crying, falling down, crawling, walking, or running away
or	
Loss of support when the body is not set for it	

Rage. The reaction is even simpler:

Stimulus	Response
Hampering of bodily movements, as head held between hands, arms pressed at sides, or legs held together	Stiffening of whole body, active movements of hands, arms, and legs, and holding of breath

Love. This term is used in its widest connotation. It includes all overt manifestations of a state of well-being and contentment.

Stimulus	Response
Stroking of skin, tickling, gentle rocking, patting	Gurgling and later violent movements of arms and trunk accompanied by laughter

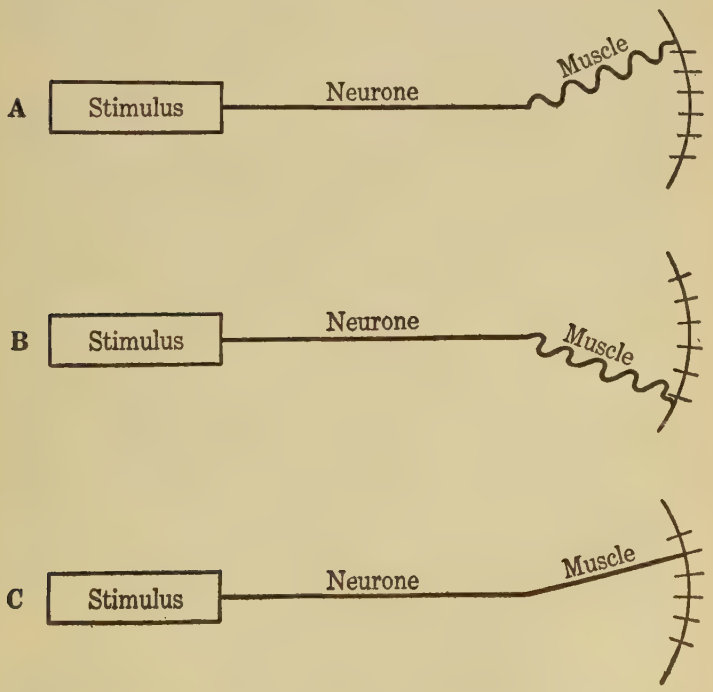
These three types of reaction, with their appropriate stimuli, constitute the groundwork of emotional life. We shall see how their range is extended to include what James calls both the "coarse" and "subtle" emotions. How the unlearned reaction of instincts and emotions is transmuted into learned reaction (acquired equipment) will be inquired into next.

B. Acquired Equipment

(1) *Learned Reactions.* We have seen that man is provided at birth with a complicated and diverse set of mechanisms which are energized by definite sources of power: instincts and emotions. We shall now examine the manner in which the mechanism is rendered more efficient and the sources of energy multiplied.

1. Neutral Coördination of Movements. This process has been demonstrated experimentally by Dr. Rivers and Dr. Head. The nerves in Dr. Head's hand were cut. No stimulus could be conveyed through it. As the nerves became welded together and the stimulus-response circuit was restored, the response was random, indefinite, and unlocalized. It was a " mass reaction " : the entire reacting organ was set vibrating. Persons under great strain react in this impotent manner. Furthermore, when this random movement abated somewhat, reaction was characterized by exaggeration : the all-or-none principle of violent emotionalism. This twofold response is known as protopathic action. When the severed nerves had regained their normal

functioning, both " mass " reacting and " all-or-none " responding disappeared. The response gradually took on a more specific adaptation. The goal was reached with less diffused excitation and with more graduated reaction. This is known as epicritic action. A diagram may prove helpful. Protopathic action would be shown thus:



The arc of maximum responsiveness, in seven units, represents the entire normal reactive range to a given stimulus. *A* and *B* show the upper and lower limit of

the responsive range reacting to a given stimulus. *A* represents the random type — the wavy line shows indecisiveness — and its position near the upper level of reactivity bears witness to its "none" type. *B* exhibits the same random movement with a maximum sweep of activity along the entire arc of responsiveness. *C* represents the decisive and direct response, with an intensity proportionate to the strength of the stimulus: in this instance two units out of a possible seven.

Epicritic movements form the raw material from which instinctive and emotional behavior is built. When these become completely fashioned, we have a highly articulated instinctive and emotional machine. But it is geared for low speed; it can execute only certain definite movements and can operate only within a narrow range of activity. It attains greater complexity through the interposition of an adaptive instrument known as the conditioned reflex. The conditioned response is a bit of learned behavior superimposed upon the unlearned response. Its mechanism is simple. Watson has found that certain stimuli bring about characteristic reactions which are constant. Such reactions are designated as standard reactions, fundamental reactions, or biologic reactions.

In the classic experiment of Pavlow the tongue of a dog was stimulated by food or acid. The salivary glands liberated a quantity of saliva. The procedure may be indicated thus:

Stimulus (S) — Food Reaction (R) — Flow of saliva

It was ascertained experimentally that only two stimuli (food and acid) were reacted to in this specific manner. We have, then, a standard stimulus followed by an equally standard reaction.

The conditioned reflex was now built in. Various stimuli of the widest sense appeal were introduced: colored discs, noises, pure tones, and bodily contacts. None of these produced the slightest excitation of the salivary glands. To none of these did the dog react with a discharge of saliva. When the experimenter applied any of the above stimuli simultaneously with the feeding, a stage of association was soon reached in which such stimuli in themselves affected the salivary glands. The new formula reads:

Powdered meat (S1)	Flow of saliva (R)
Red disc (S2)	Flow of saliva (R)

We might express this formula differently:

$$S1 = R1$$
$$S1 + S2 = R1 \text{ after many repetitions}$$
$$S2 = R1$$

Then,

$$S3 = R1$$
$$S4 = R1, \text{ etc.}$$

" Conditioning " is the term applied to this substitution of stimuli. Infants have been conditioned. The sight of a pipette has no effect. When the pipette is filled with acid and applied to the tongue, the stimulat-

ing quality of the acid is transferred to it. Again, muscular activity has been conditioned. Watson * works out a formula like this:

Cutting, bruising, electric shock (S)	Withdrawal of arm, leg, or finger (R)
Closely followed by sound of electric buzzer (auditory stimulus)	Same response

There is reason to believe the same conditioning takes place in the internal organs as well. They cannot be adequately accounted for by the psychological process of association of ideas. The infant does not detect the likenesses; his associative powers are weak. Again, these glands are not in the orbit of voluntary control. They belong to the sympathetic (autonomous) nervous system. This is truly a building in by environment and not an expansion from within.

The experimenter exercises conscious direction. The environing world acts fortuitously. Numberless stimuli become integrated through repetition with a standard stimulus, any one of which may provoke the concomitant reaction. The individual may be ignorant of this organized unit of behavior. Let one of these substitutive stimuli present itself and the individual reacts characteristically. This bit of behavior may be wholly inconsistent with the setting in dominant behavior, and

*Watson, John B. — *Behaviorism*, W. W. Norton and Company, 1925.

he may manifest as much surprise at this aberrant behavior as his friends. The returned soldier who threw himself upon the ground upon hearing a tire explode was no less distressed upon recovery than is the individual who in the quiet of his home jumps at the sound of a horn.

2. The Conditioned Emotional Response. The original emotional equipment undergoes a similar transformation. The emotions become diversified and develop a wide range of unit responses. Conditioning follows the well-known pattern of stimulus-substitution-response.

Sight of white rat (S1)	Manipulation (R)
Loud sound (S2)	Fear (R)
Sight of white rat (S1)	Fear (R)

When infants are confronted with animals, they react positively and pleasurably. Let the experimenter produce a loud noise at the moment the child grasps the animal and he exhibits all the signs of intense fear. Rage and love change rapidly from a state of undifferentiated behavior into one of highly differentiated response. The latter is attained through the agency of another diversifying mechanism, called transference.

The child who becomes conditioned (exhibits the emotion of fear) to a rat will react similarly to other furry animals or even furry objects without further conditioning. He will be thrown into a paroxysm of fear in the presence of a rabbit, a dog, a pigeon, human hair,

cotton wool, a fur cap, etc. However, further conditioning may be carried on to the point of habituating the responding organ to selective responses: certain stimuli possessing a range of intensity may organize differential responses. Strike the tone A while feeding the baby and conditioned response will at first switch to that stimulus and its variations. Continue this practice and feed only when A is sounded — not when other notes are struck — and you soon form the tone-A-desire-for-food circuit. Tone A becomes the unique stimulus for this response.

3. The Significance of Conditioned Reactions. The bodily sets that the infant inherits are few. We have seen that the number of instincts and emotions are small. They soon gain in flexibility and variety largely through chance conditioning-organic association, if you will. Differentiation and transference enlarge the repertory of responses, and the unlearned reactions become overlaid with a thick layer of learned behavior. Since the learning process becomes operative at birth, psychologists who study human behavior in early childhood and adulthood are led to regard all autonomous and habitual responses as the contribution of heredity. They are further misled by the absence of conscious direction. They know how associative bonds are formed in the schoolroom, and they view them from a purely intellectual angle: the summation of perception, imagination, conception, memory, judgment, and reasoning. They do not ignore feelings and

emotions; they assign them subsidiary rôles. When confronted with manual and bodily responses, they recall the feebleness of infantile mentality and promptly classify them as innate tendencies. Whence the designation of a multiplicity of instincts and emotions and the failure to appraise them at their true value.

For an adequate evaluation of bodily sets and attitudes we must go to the Behaviorists. They envisage personality as segmental in its development, each segment constructed of an infinite number of units (reactions) organized around a fundamental stimulus and its accompanying response and involving profound bodily changes. They further maintain that these changes are brought about in infancy at least without the direct intervention of the central nervous system; that the environment exercises a tyrannical domination, modifying, altering, building up, and transferring arbitrarily and capriciously; and, most importantly, that these chance associations color every act, both emotional and mental, of every individual. Take any stimulus-response situation, such as learning $5 \times 6 = 30$. It may become organized into various units of reaction in which S = visual stimulus ($5 \times 6 = 30$) and R = laryngeal movement (repetition).

(1) S = R

Now suppose that the child is indisposed at the moment of making the reaction. Or the teacher may have scolded him for misbehaving directly after the stimulus was presented. Then you would have:

(2) S = R + V (visceral accompaniment — unsatisfyingness)

Again, the numbers 5 and 6 may have become associated with a painful experience: the child may have been late several times and duly warned or threatened with dire punishment. The reaction would advance through two phases:

(3) 5 × 6 (S1) = 30 (R1)

5 (times late — S2) = Fear of punishment (visceral accompaniment — R2)

5 × 6 (S1) = R1 + V

Or take a manual circuit: learning to write.

(1) Sight of letter (S) = Copying (R)

The teacher may have rapped the child's knuckles for holding the pen in a cramped position. We get:

(2) Bruising knuckles (S2) = RV2 (withdrawal of hand, and rage)

(3) Sight of pen (S1) = R1 + RV2 (visceral accompaniment)

We shall revert to this phenomenon from time to time to impress the fact that feeling tones tinge our behavior, that the effect is either to advance or retard the response, and that there is no such thing as pure intellectual activity. Every mental act is bathed in feeling arising out of visceral and vasomotor changes.

4. The Conditioned Reflex and Habit. The conditioned reflex, with its threefold constituents — the fortuitous or conscious stimulus, the biologically

adequate stimulus, and the specific response — is closely akin to habit. It differs from habit in two particulars: the depth of the concomitant emotional tone and the close association with the fundamental stimulus. But the organism also acquires responses that are not associated with the biologically adequate stimulus. They arise from the activity of the stimulus-response mechanism — and become fixed through repetition or shock, as a child burning his fingers through contact with a candle or match. A pathway has been made which will persist indefinitely with disconcerting tenacity. This is an example of the unconditioned reflex through shock. Mere repetition may take the place of " shock." The manner of holding a spoon is built into the behavior by repetition assisted by satisfiers and annoyers.

Both conditioned and unconditioned reflexes may be regarded as habits if the reader will keep in mind the emotional connotation invoked by the other terms. Max Meyer hits off the distinction between conditioned reflex and habit rather neatly. " If any student," he writes, " has the habit of thinking of a habit as a sheer mystery, then he should be advised to clarify his thought by using the term ' conditioned reflex,' which would constantly remind him of the fact that no habit can be put on an animal except one whose sense function and motor function have already, in other combinations, been given to the animal by nature." He who realizes this fact, that " habits " cannot be created out

of nothing but only of "reflexes," need not use the inconveniently long expression "conditioned reflex."

Habits, then, are organized reflexes. In time they form complex units in which elements of simple habits are reassociated, forming a hierarchy or habit system. The learning process is nothing more than a hierarchy of habits: specific habits yield common elements which are combined in the intervals of complex learning progression known as plateaus or resting places.

5. Attitudes. We have seen that responses, both elementary and complex, which constitute behavior are bathed in feeling; that there is no response which at some stage of its existence does not tap the emotional reservoirs of the organism. When this feeling tone is intense and subsequent analysis and rationalization have not purged it of fear elements, we have a complex. In their subacute state, they color the response with which they are linked, into an attitude. Since habits are responses, every habit has an attitude correlated with it. Foremost among them are the generic attitudes toward life, such as depreciation, conceit, aggressiveness, indifference, optimism, pessimism, antagonism, helpfulness, enthusiasm, depression, and courage. Wrong attitudes are, according to Abbot, sufficient to produce *dementia præcox*. "If," he succinctly states,* "a child keeps by itself; does not seem able to get on common ground with other children, gets absorbed in

* *Preventable Forms of Mental Hygiene*, Massachusetts Society for Mental Hygiene.

its own daydreams and fancies, and resents being called out of them; is sensitive and takes knocks and disappointments as slurs or injustices, resents them, broods over them, and regards itself as a martyr; nurses its grievances in its own breast, not confiding its troubles to others but continuing the sense of injury and injustice; in general, keeps its thoughts to itself; and when its chosen occupation is denied, goes off by itself with a sense of injury and does not try to take up another, but lets the mind dwell on unproductive fancies and vague, impractical aspirations — such a child is more likely than the others to have *dementia præcox*."

Besides these generic attitudes there are innumerable specific attitudes conditioning all school work. These come into being through adverse environment or through the teacher's reactions to both pupil and subjects taught; and it is hardly an exaggeration to say that half the energy expended by both pupil and teacher are frittered away in the attempt to counteract wrongful attitudes. It reminds one of the type of problem, no longer admitted into textbooks, which asks, if one pipe fills a tank in three hours and the other empties it in two hours, how long it will take to fill the tank. The answer is obvious no less for the illustrative problem than for the psychologic one of stopping this emotional waste; no precept, no matter how practical, can overcome an emotionally tinged propensity. No pedagogic competence can avail against the I-hate-arithmetic attitude.

In its rough outlines, good teaching results in self-active learning applied to such innate endowments of the individual as sense organs of adequate acuity, capacities, and the degree of intelligence. Normal forgetfulness takes its toll of learning. This is the chief subtractive factor. But when an obnoxious attitude whose core is unsatisfyingness supervenes, the net scholastic attainments are considerably decreased. The situation confronting the teacher is discouraging, for we have the following conditions which should be considered:

Teaching process = Self-active learning process + Hereditary endowments (sensorium, capacities, intelligence) − Forgetfulness

Teaching process = Self-active learning process + Heredity − (Forgetfulness + Resistant obnoxious attitude)

(3) *The Learning Process*

1. Inadequacy of the Trial-and-Error Principle. Trial-and-error learning, with its basal S–R process, describes a stage of behavior at the lowest level of intelligence. The principle of trial and error or success and failure has been elaborated as a result of the experimental investigations of Thorndike performed upon cats and dogs confined to a cage from which they sought to free themselves in order to reach food within plain view. The cage was provided with a door which was closed by a simple mechanical device operated by pull-

ing down a string, turning a lock, or pressing upon a board. The investigator reports upon the behavior of the animal from the time it was introduced into the cage to its final liberation. Stress is laid upon the random movements of the animal, the time consumed in opening the door, and the number of errors made during the process of springing the lock.

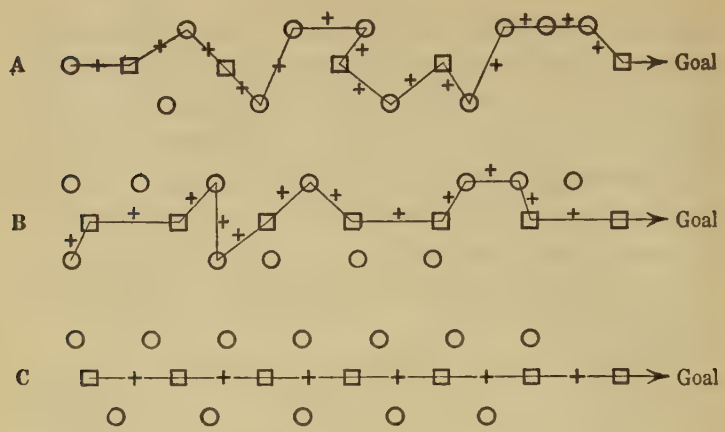
The behavior of the animal is ascribed to the following considerations:

1. The random movements involved in the gross attack upon the bars of the cage are initiated by the instinctive repertory of reactions. Thus the hitting out at random, the grasping of the bars follow the S–R mode of activity.

2. The random movements continue, the successful and unsuccessful ones mingled as one unresolvable confusion.

3. The reactions continue intermittently in this inchoate state and are terminated accidentally by a chance movement which proves effective; the door is flung open and the animal makes its escape.

4. Upon many repetitions of this experience, the animal gradually effects a reduction of the random movements to the point of eliminating all useless motions and of linking the useful movements in a series in which the specific S–R reactions are closely associated by practice. If ○ represents useless movements and □ useful movements, then the learning by trial and error proceeds in this wise:



5. The animal is not an active participant in this learning process. The modification of its behavior takes place without the slightest notion of the import of the proceedings. Hence the length of the time curve and the unbelievable stupidity of the errors it commits. Understanding follows only upon the completion of the practiced successful movement.

2. The Rôle of Purposiveness. The trial-and-error principle suffers from many defects, chief of which are the passivity of the learner and the assumption that the S–R formula is sufficient to explain all forms and stages of learning. We shall deal with these successively in detail.

We have seen that chance stimuli may so insinuate themselves into the total personality as to effect enduring changes in behavior. Others may achieve a factitious interest by virtue of the emotional complex which

may be built around it. There are, then, many occasions when the personality becomes submerged and subjugated by vagrant stimuli that assume dictatorial powers.

But life, neither for man nor for higher forms of animal life, is a rational or, at times, irrational linkage of environmental and instinctive forces welded together by association, for there is a power above the more or less compounded mechanical reflexes, a power that derives its potency from the free-floating energy back of the instinctive " drives." It manifests a latency that carries the organism far beyond the boundaries of objective stimuli into the more trackless sphere of imaginatively conceived modes of behavior — behavior that is untrammeled by exigencies of the stimulus-response situation and its constant menace to self-determination.

This glimpsing into the world beyond the senses, this living on the plane of the subjective, this thrusting beyond the physical, this restless exploration of the future is called purposiveness. It may be defined as the impulsive urge of the personality to desire and strive for the attainment of a goal, mediated by a number of discrete S–R links all of which may not be fitted into a chain but which possess much distractibility and untold possibility of deflecting the aim.

Purposiveness has been variously defined as striving, motivation, incentive, struggle for satisfaction, craving, aim, means to an end, and autonomic drive. They all seek to make clear that the organism strives to partici-

pate in all attempts to modify its behavior when the stimulus or response is deprived of the driving power of shock or the fixating power of prolonged visceral stress. Or, as McDougall phrases it, " *Purposive* is the adjective which we may use if we wish to describe objectively processes of a special kind, namely, the peculiar forms of bodily action we observe in men and animals, the goal-seeking actions. But we may use it, as we do in common speech, to imply also that such actions are accompanied by some foresight of the goal and some desire to attain it."

3. Total Situation *versus* Compounding of Its Elements. We shall now inquire into the inadequacy of the S–R formula as a complete description of the purposive learning process. The stimulus-response hypothesis describes the act of learning as an accretion of specific S–R elements held together by association. It assumes that these elements are amenable to the law of practice; that, obedient to the law of association, the perception of one of the elements reinstates the movement-complex in which the newly acquired behavior consummates; and that this newly acquired behavior arises out of the matrix of instinctive acts with which the organism is endowed.

In order to confute the theory, we shall study the classic experiment of Köhler on anthropoid apes. He placed a number of apes singly in cages and placed bits of food within full view but out of reach of the inmates. Inside the cage he introduced in succession sticks and

boxes with a view to testing the inventiveness of the apes in reaching the food, through the intermediation of objects employed as " tools." Thus the direct way to the goal was barred.

The chance solution resorted to by Thorndike's animals would lead us to expect a random handling of the objects within the animal's field, leading to an accidental grasping of the stick or the chance placing of the box so as to facilitate the attainment of the goal — the seizure of the food. What the apes did was something entirely different. A period of random reaching after the food alternated with an equally meaningless handling of the stick or box. Then there ensued a period of quiescence, a taking stock of the situation. The end came swiftly and unhesitatingly; the ape grasped the stick and knocked down the food, just as in another experiment it laid hold of the box and, placing it under the food, stood up and seized it. Thereafter the act was performed in one continuous movement.

True, we must concede a period of trial and error. But it did not dominate the behavior of the animal and was soon discarded for a more deliberative mode of activity. This sudden emergence of the successful behavior is explicable on grounds other than that of chance-activity favoring the apes. The solution was discovered rapidly because intelligent learning, as evidenced by the experiments, employed the following technique:

1. The goal, in this case the seizure of food, was kept steadily is view.

2. The objective elements in the field of vision were neither manipulated blindly nor were they regarded as arbitrary objects, the preliminary random handling of them notwithstanding.

3. During the cessation of random handling of the objects, ascribable in greater part to the eagerness evinced in compassing the end, the behavior of the animal was motivated by the exercise of insight, the awareness of the interrelationship subsisting between the intervening objects and the goal. Understanding, it would seem, preceded the successful behavior.

4. Insight operated to aid the animal in envisaging the discrete elements as members of a total situation, as the Gestaltists have it, as members of a coherent, sequential configuration.

5. This flash of insight was followed by achievement, the reaching of the goal through a continuous curve of progress extending from the point of departure to the consummation of the movement-complex.

6. The successful mode of behavior was fixed at the moment of achievement and, unlike trial-and-error learning, did not exhibit, upon repetition, as large a curve of errors.

A true solution of a behavior problem proceeds as follows:

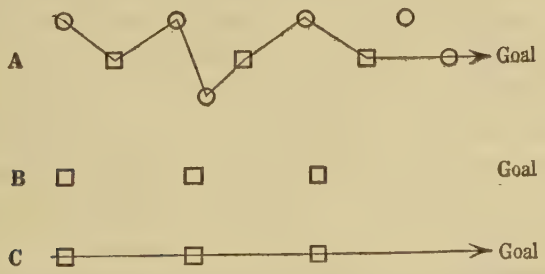
1. A highly transitory period of trial-and-error behavior initiates the learning process.

2. The goal-seeking proclivity asserts itself in the calm contemplation of the sensory field.

3. By the aid of insight the disconnected stimuli are apprehended as members of a pattern or configuration. Only those stimuli are accorded membership in the configuration which impart meaning to the movement-complex.

4. The elements of configuration become, then, as durably fused as notes in a melody, so that moderate repetition tends to preserve the alignment.

5. Employing the symbols ○ and □ to represent respectively unsuccessful and successful movements, or the handling of right objects, the true solution may be shown thus:



A. Random Handling of Objects, without Attainment of Goal
B. Alignment of Objects or Contemplation of Movements
C. The Stage of Achievement

6. According to the Gestaltists, configural learning marks a stage of original acquisition in that the organism does not evolve the newly acquired behavior from innate dispositions brought to the focus of consciousness by the perception of the sensory elements of

the new situation. The fruit and the stick in Köhler's experiments are not brought together because, as Buhler maintains, " an arboreal animal must be quite familiar with the connection of branch and fruit," and therefore the conjuncture of a " branchless fruit" and a " fruitless branch " (the stick) leads it " to incorporate the two, so to speak, in one idea." In other words, chance discovery, motivated by instinct and bound together by practice, is sufficient to explain even indirect learning or learning by the creation of " tools."

But the Gestaltists refuse to place such an interpretation upon the animal's use of the stick. For the stick is manipulated as a " stick " until the moment insight places it within the configuration of " stick-tool-fruit." This consideration leads Köhler to decide that there exist no " opposed relations of a fruit without a branch and a stick . . . before they are united; on the contrary, all the perceptual situation contains at first is a desired but unattainable fruit (goal seeking) together with a quite irrelevant stick "; and that the complex stick-tool-fruit is a totally new configuration arising less from instinctive and ideational (memory of branch plus fruit) matrix than from a sudden awareness of the true inwardness of the situation. The field of perception, then, is not welded together by sheer manipulative force applied at random at the behest of an instinctive urge, but is molded into a configurative whole by the exigencies of goal seeking operating through relevant objects in the total situation.

4. Types of Learning. These types may be divided into four classes.

(*a*) *Learning by Accretion.* This type may best be illustrated by Yerkes's experiments with the earthworm. Into the stem of a T-shaped maze, an earthworm is introduced. After crawling through the stem bar, the animal may turn to the right and reach its burrow or it may turn to the left and receive a painful electric shock. It learns to avoid the left turn in 20 to 200 trials and in a reduced number of trials to make a right turn when conditions are reversed.

The method employed by the earthworm is the forced association of movements involving a low grade of intelligence and little conscious participation. It is not confined solely to animals; human beings make use of it when spurred on by fear or driven to it because of lack of insight or of inability to perceive a configuration in an incongruent assemblage of perceptual elements. The school offers many examples of such inept learning, which is inherently artificial. Ogden aptly characterized this educational method as " putting pupils through the act of learning." Its chief reliance is the accretion of facts or ideas through forced association, an application of trial-and-error learning.

(*b*) *Learning by Assimilation.* The chimpanzee of Gestalt fame exemplifies the assimilative type of learning. When placed in a cage with a banana and a stick or, as a variation of the experiment, two sticks that may be fitted into each other fishing-rod wise, he reacts

for a time violently and blindly to those elements of the perceptive field which become the objectives for the time being. He reaches vainly for the fruit, shakes the bars, grasps the stick, pokes with it at the bars or floor. Should chance manipulation bring the stick in contact with the fruit and force it down, there is no evidence that the chance contact will be repeated or, if repeated, that the act came about by design. An alteration in the animal's behavior, does it occur the moment the stick and the banana fall together in the line of regard, and insight combine the two into one pattern; then the stick is seized and employed as a tool; or, later, when the banana can be made accessible by the simultaneous use of two sticks, the two are fitted together and the problem is solved.

This is assimilative learning because the animal has effected a synthesis between the perception of the banana and the perception of the relevancy of the stick; he did not merely add to the perception of the banana the perception of the stick as a stick which chance movement caused to release the banana. Likewise larger syntheses are formed, not by accretion of selected factors but by the sudden and unlabored perception of those congruent elements which merge patternwise.

(*c*) *Learning by Gradation.* Once more are we indebted to Köhler, who furnished convincing data derived from experiments performed on hens and repeated on apes and a three-year-old child. He employed two gray papers of different shades, each hold-

ing an equal number of kernels. These were placed on boards before the hen coop. The hen was permitted to take the kernels from only one of the papers, the lighter of the two. In order to preclude the possibility of choice by position, the papers were interchanged at intervals. Some hens learned to make the right choice in 400 trials; others, in as many as 600. In the control test a lighter shade was substituted for the favored one. The hen promptly chose the lighter of the two papers.

The response was striking in that the dark shade, which remained constant in both experiments, failed to have associated with it the same unequivocal response indicated by the law of practice. For here we have all conditions favorable to trial-and-error learning — random movements, chance discovery of the acceptable response, the application of rewards and punishment, then chance attainment of a goal, and the gradual forced association of the successful movement with it. Yet the hen chose the response to the stimulus which bore the relation sustained by the papers in the original training tests: the *lighter* of the two shades was preferred.

What seems to have influenced the hen in the choice of the shade was the degree of gradation obtaining between the two. The hen recognized the presence of a gradient, light to darker, and made an immediate successful transference from the light shade of the training tests to the lighter shade of the control test, obedient to the altered relation in the new arrangement of

the papers. She reacted to a gradient of brightness-darkness that manifests "a tendency toward" from one to the other; a directional pattern emerges from this rhythmic relation of stimuli possessing the coherence of a configuration.

The search for the gradient which inheres in stimuli, whether objects or symbols, percepts or ideas, constitutes a third type of learning, having the agglutinative property of a configuration and the characteristics of a higher stage of intelligence, intelligence calling for a more subtle discrimination of relationships. It is through the perception of gradation that the child learns quantities of more or less and the nuances of language.

(*d*) *Learning by Redefinition.* We have seen that a field of perception remains static and neutral as long as the stimuli are not woven into a contest; the endowment of learning transforms them into a configuration; they become figured or structured. In configural learning, meaning is invoked whenever the total situation suffers a modification or alteration. Since the total situation varies with reorientation of the goal, meaning undergoes a corresponding change. The configuration

formed from four slightly curved lines may appear as a shield, a sail, or a square with curved sides. By an act of recentering or refocusing, redefinition has been achieved.

Redefinition is a highly educative process when responses are employed with ever varying meanings, depending on the richness of the fund of experience which the child brings to bear upon old situations. Especially is this type of learning potent in what Ogden calls " equivocal figures of apprehension," situations pregnant with varied meanings any one of which may leap into consciousness through an act of refocusing.

5. Summary of the Learning Process. The various phases of the learning process might be summarized thus:

All learning arises from the manipulation of the stimulus-response situation. Both stimulus and response may be conditioned by the process of " building " them in both singly and in series as habits and attitudes. The total personality participates in all acts of learning. The participation is sometimes passive but more often active and selective, depending upon the keenness of the need felt for adjustment and in keeping with the goal-seeking proclivity displayed by all higher forms of life. The quest for self-realization is its highest manifestation; and, when thus employed, objects in the immediate field of perception are manipulated both as neutral independent stimuli and, in the event of their forming an easily discernible pattern, as means of gratifying a desire. The discovery of the pattern serves two purposes, that of establishing a direct route to the objective and that of gaining memoriter aid through the emergence of a configuration. The

former gratifies a felt need; the latter makes future gratification more facile.

The cohesive qualities of configurative learning are productive of equal facilitation in solving problems of personality adjustment in situations in which the constitutive stimuli are organized in more subtle ways. Objects in the environment may be related to each other in the form of scales, rising or falling as gradients. Here the direction of change, as in difference of size or constantly varying geometrical designs, indicates the presence of a configuration.

Equally subtle is that type of learning which infuses new meanings into a contest as a result of the reintegration of the elements. By changing the focus of attention and by modifying the mode of orientation, old configurations are resolved into newer ones, thereby adding flexibility to the power of personality adjustment. The designation "redefinition" has been applied to this type of learning.

Finally, the personality achieves greater efficiency in self-determination by the method known as transfer of training. By the transfer of training is meant the ability to solve new problems of adjustment by means of terms learned in solutions involving similar conditions. A learned response evoked by the interplay of insight and purposiveness directed toward the connected elements of a situation tends to function readily, under conditions in which these elements may either become reassociated into new configurations or may

coalesce with similar elements of allied situations, without the expenditure of as much mental and physical energy as was exerted in the mastery of the elementary situation. Much of the learning thus acquired employs the economics of transfer of training.

These principles of learning owe their elucidation to the investigations of Köhler and his newly arisen school of Gestaltists, who have labored in the field of animal psychology. They have enunciated clearly the principle of configurative learning as a result of experiments conducted with anthropoid apes. They owe their success to the improved technique of experimentation — a technique which recognizes the imperativeness (1) of creating for the animals an atmosphere of naturalness, (2) of not exceeding their discriminatory powers, (3) of affording them situations which accelerate the construction of configurations, and (4) of employing a minimum of mechanical devices. Thorndike and his school of investigators erred in placing the animals in boxes — an environment unnatural for dogs and cats — in tantalizing them with locks and strings which they had not learned to associate with food getting, and in obscuring the configural relations between these mechanical devices and the seizure of the food. Small wonder, then, that their reactions were both stupid and unproductive.

C. Directive Equipment

The discussion of Köhler's experiments upon apes, with its emphasis upon the display of insight antecedent

to the configural transformation of the perceptual field, may have exaggerated the rôle played by intelligence. Even the elimination of Thorndike's chance-hypothesis and its chief supporting arguments — (1) the form of the time curve and the contention that time measurements of animals' reactions point to chance-insight and (2) the abundance of " stupid " errors — but leads Köhler to favor the theory of the chance-situation becoming patterned when the relations between the perceptual field and the goal grow " perspicuous." True, the experimental data he gathered incline him to the belief that solutions in which insight plays so important a part merit the characterization intelligent. But more complicated experiments have forced upon him a cautious attitude with respect to the degree of intelligence exhibited in configural learning. In his so-called " detour-board " experiment and its elaborations, the animal sat near the bars facing a drawer with open top and lacking a rear wall. Food was placed in the drawer nearest the animal, who was given a stick with which (1) to free the food in the drawer by pushing it away from him to the open end or (2) to push it sideways. " Only when it was completely outside the drawer could it be brought forward." The animals achieved but indifferent success because the solution of the problem required a high degree of intelligence, for chance that brings the perceptual field in alignment with the goal (chance-situation) and insight which grasps their significance were not closely implicated. Their relations

were obscured by intermediate ends and the solution retarded by the contradictory tendency, that of pushing the food away from the animal instead of, as in the " tool " experiments, bringing it forward.

The quality of intelligence involved in configural learning, while decidedly higher than that exhibited in trial-and-error learning, proves unequal to a task complicated by detours, primary and secondary goals, and contradictory tendencies. When the latter conditions are multiplied so that sound judgment must be exercised in recognizing the consecutiveness of intermediate steps from stimulus to goal and tendencies must be inhibited that run counter to instinctive trends, we may say that such problem solving requires an act of intelligence of a higher level. All levels above the one which operates in the attainment of goals set in a simple context whose meaning may be penetrated by insight have for their essential features :

(1) Secondary goals (getting the food out of the drawer in the " detour board ")

(2) An involved context requiring clearness of apprehension

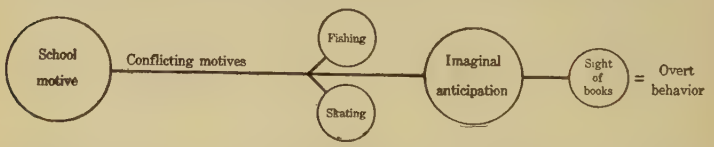
(3) Stimuli of less relevancy to the total situation (the stick and the indirect use it has to be put to to make the " detour ")

(4) The holding in check of a natural response (the tendency to bring the food forward in a direct movement)

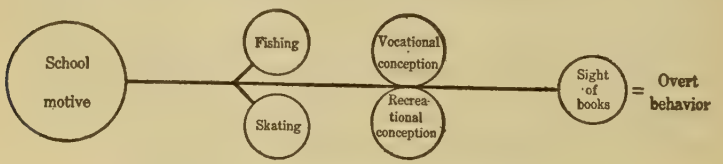
We shall now study the variety of responses possible

in the simple act of going to school. If the school motive is strong, the mere sight of schoolbooks is sufficient to initiate the act. The behavior occurs on the plane of perception.

Let us now assume that the inner urge to go to school is not strong. This mode of reaction may take place:



Imagination aids in resolving the conflict; it warns the individual against the truancy motive, drawing upon memory for corroborative evidence and pointing out to the vacillating mind that direful results might follow. Again, imaginal anticipation might prove ineffective. Reaction might be delayed at the conceptual stage of thinking; the benefits of schooling might be compared with the joys of untrammeled activity, with fear pitted against the educational motive, the reasoned advantages of good attendance against the fugitive pleasures of recreating, in this way:



We have noted that the exercise of intelligence has the effect of slowing up the motor tendency of motives, desires, and impulses so as to permit relating them conceptually and imaginally to be what is best in the individual's environment. Furthermore, intelligence weighs and evaluates conflicting motives and renders the decision in the form of a judgment. The appropriate percept (or image) calculated to bring about the satisfying behavior is then sought out and made focal in consciousness. You will recall how the sight of the books precipitated the reaction (going to school).

Intelligence also accumulates the experiences of the past and present, organizes them into systems, both accretive and configural, and makes them available in deliberation. Behavior ceases to be impulsive; action is delayed to the end that obnoxious responses be inhibited and favorable responses be enhanced. In the light of its directing and appraising function, then, intelligence may be defined as the psychic activity which stores up the experiences of the past and the present so that they can be used in the future, and adds to the rigid mechanism of instinct the flexible mechanism of reasoning. This is not a psychological or pedagogical definition. It is purely a biological explanation in terms of coördination and adjustment.

Personality

We are now in a position to gather together all the threads of the discussion and unify them in a compre-

hensive definition of personality. We may define personality as the psycho-physiological organization of innate and acquired dispositions, tendencies, and traits, by virtue of which the individual is enabled to harmonize the dissident forces of instincts and ideals — such innate dispositions and tendencies as manifest themselves through instinctive and emotional behavior and such acquired traits as sentiments, beliefs, habits, mental attitudes, and mental sets. Warren follows the traditional view of personality and defines it as " the entire organization of a human being at any stage of his development. It includes all four phases of human character — intellect, temperament, skill, and morality — and every attitude that has been built up in the course of one's life." Warren is right in the fundamentals, though the arrangement is faulty. Genetically, temperament comes first, then attitudes, followed in order by intellect, skill, and morality. Again, attitudes might be either mental or emotional and are implicit in the development of intellect and temperament. Skill, which is both manual and intellectual, may be subsumed under " intellect." Warren's definition may be reduced to a three-dimensional measure of personality: (1) temperament, (2) intellect, and (3) morality.

However, there is still another dimension: it is predominantly motor and concerns itself with the strength of inner drives, the urgency of impulse, and the propulsive power of psychic and physical energy. Myerson classifies types on the basis of the force and volume of

the discharge of energy and recognizes the hyperkinetic (the high-tensioned energy discharger), the mesokinetic (medium or average energy discharger), and the hypokinetic (feeble energy discharger). A fourth dimension must be added, the kinetic. Personality, therefore, may express itself in terms of temperament, intelligence, morality, and activity. Activity may, with good reason, be placed beside temperament because the discharge of energy may be contingent upon the absence or presence of irascibility, which in turn determines the disposition of the intellectual and moral forces.

Temperament, which has regard to the emotional equipment of the individual, is based in part upon glandular harmony or lack of harmony. The complex also serves as a determinant. Three grades may be differentiated: (1) the cyclothymic (the irascible) — the overemotional, with its chief variant, the monothymic (with a dominating emotionism), (2) the balanced, and (3) the epileptic. The classification might be rendered more specific by indicating the presence or absence of complexes. An individual might be cyclothymic, with few complexes, or cyclothymic, complex-ridden. The same modification might apply to the balanced and the epileptic. We shall discuss tests for uncovering complexes in another chapter.

Intelligence will be discussed fully in subsequent chapters.

It should have become apparent that the current

measures of intelligence are hopelessly inadequate in measuring the entire child. True, the intelligence test is a fairly efficient tool for measuring intelligence, but the intelligence test is no more diagnostic of pedagogic efficiency than the thermometer is of disease. Learning is a unitary act in which the entire personality participates. Scholastic progress may be retarded by conflicting " non-intelligence " traits, such as emotional excitability, the type of energy discharge, the acuity of the sense organs, the degree of cogency displayed in reasoning, the retentiveness of memory, and to a lesser extent the standards of morality cherished by the individual. Morality is a social requisite in that knowledge may have a social or antisocial application to the unsettling of social stability.

Regarding the kinetic dimension, we shall follow Myerson in his classification of types of energy discharger. This category may include attributes often included under moral qualities, such as industry, applicability, the permanence or impermanence of interests, and perseverance under adverse conditions. We might classify a child as mesokinetic, applying himself but moderately or intermittently to a task. It is this very attribute (applicability) that renders intelligence scores significant. The pupil with the high I.Q. who does not apply himself to his studies is the despair of both teacher and mental tester. Yet another subclassification may be desirable. This kinetic dimension may include the predominantly active and the

predominantly thinking type. Jung has named them extraverted and introverted respectively. Introversion means self-absorption, a seeking to gratify desires and wishes in internal readjustment, such as an alignment of attitudes or an elaboration of the dream life ranging from the phantasy to delusion. The extraverted type is absorbed in the world without. From these come the "luncheon-clubbing, sport-dominated kind," who in childhood stressed doing rather than learning, manipulation rather than abstract thinking, in short, the "doers" and not the "dreamers." Indeed, this dimension might as well be expressed as introversion, ambiversion (a combination of the two tendencies), and extraversion.

Exercises for Self-Testing and Investigation

I. Indicate whether the following statements are true or false.

1. Temperament is inherited.
2. Endocrine glands influence personality.
3. The thyroid glands control growth.
4. The pituitary gland is located in the throat.
5. The adrenals influence emotional life.
6. Emotions are acquired.
7. The term "conditioned reflex" was introduced by Thorndike.
8. Personality is predominantly intellectual.
9. Instincts are unlearned responses.
10. Attitudes are unlearned.
11. The cerebrum controls thinking.
12. The quality of memory is conditioned by the medulla.
13. The mechanism of stimulus-response controls all learning.
14. Consciousness is a form of chemical reaction.

15. Configuration is a concept of Gestaltism.
16. Trial and error is a form of intelligent learning.
17. Intelligence controls behavior.
18. Complexes result from suppression.
19. Morality is a phase of personality.
20. Ideals are ideas held emotionally.
21. Anger is an instinct.
22. Psychologists are in disagreement as to the number of original instincts.
23. The individual learns with his entire personality.
24. Reflexes are instincts.
25. Reflective thinking is acquired.

II. The following psychologists are known for the following contributions. Indicate whether (*a*), (*b*), or (*c*).

1. Watson:
   (*a*) Study of animals
   (*b*) Conditioned reflex
   (*c*) Studies in personality
2. Freud:
   (*a*) Neurotic individual
   (*b*) Complex
   (*c*) Gestalt theory
3. Adler:
   (*a*) Individual psychology
   (*b*) Psychologic types
   (*c*) Inferiority complex
4. Koffka:
   (*a*) S–R formula
   (*b*) Gestalt theory
   (*c*) Complex
5. Thorndike:
   (*a*) Localization of brain areas
   (*b*) Trial-and-error learning
   (*c*) Studies in fatigue

6. Bergson:
   (*a*) Learning process
   (*b*) *Élan vital*
   (*c*) Rationalism

7. Rivers:
   (*a*) Psychology of reasoning
   (*b*) All-or-none response
   (*c*) Educational sociology

8. Watson:
   (*a*) Structuralism
   (*b*) Purposivism
   (*c*) Behaviorism

9. Freud:
   (*a*) The subconscious
   (*b*) The unconscious
   (*c*) The apperceptive mass

10. McDougall:
    (*a*) Purposivism
    (*b*) Psychoanalysis
    (*c*) Gestaltism

III. The following psychologists are credited with one of three utterances, one of three books, or one of three tests. Indicate whether (*a*), (*b*), or (*c*) is correct.

1. Terman's definition of intelligence:
   (*a*) "The ability to carry on the work of the class"
   (*b*) "The ability to make the adjustments to society"
   (*c*) "The ability to carry on abstract thinking"

2. Van Wagenen:
   (*a*) Reading Scales
   (*b*) Fundamental Processes in Arithmetic
   (*c*) Handwriting Scale

3. Baker:
   (*a*) Point-Scale of Intelligence
   (*b*) Stanford Revision
   (*c*) Detroit Advanced Intelligence Test
4. Myers:
   (*a*) X–O Cross-out Test
   (*b*) Mental Measure
   (*c*) Reading Scale
5. Buckingham:
   (*a*) Illinois Examination
   (*b*) Kindergarten Test
   (*c*) Revision of Binet Scale
6. Buckingham:
   (*a*) *The Measurement of Intelligence*
   (*b*) *Diagnostic Teaching*
   (*c*) *Research for Teachers*
7. Terman:
   (*a*) *The Psychology of Subnormal Children*
   (*b*) *Genetic Studies in Genius*
   (*c*) *Teaching Dull and Retarded Children*
8. Ayres:
   (*a*) Arithmetical Scale
   (*b*) Handwriting Scale
   (*c*) Reading Scale
9. Thurstone:
   (*a*) *Vocational Guidance*
   (*b*) *The Gifted Child*
   (*c*) *The Nature of Intelligence*
10. Whipple:
   (*a*) Handwriting Scale
   (*b*) Test for Gifted Children
   (*c*) Arithmetic Scale

Bibliography

Allport, F. H. and G. W. — "Personality Traits: Their Classification and Measurement," *Journal of Abnormal Psychology and Social Psychology*, 1921

Berman, L. — *The Glands Regulating Personality*, Macmillan, 1921

Brown, W. W. — "Character Traits as Factors in Intelligence Test Performance," *Archives of Psychology*, No. 65, 1923

Burks, B. S. — "A Scale of Promise," *Pedagogical Seminary*, September, 1925

Burnham, W. H. — *The Normal Mind*, Appleton, 1924

Conklin, E. S. — "Definitions of Introversion, Extraversion, and Allied Concepts," *Journal of Abnormal Psychology and Social Psychology*, Vol. 17, pp. 367–382, 1923

Downey, J. E. — *Will-Temperament and Its Testing*, World Book, 1923

Folsom, J. K. — "Statistical Study of Character," *Pedagogical Seminary*, Vol. 24, 1917

Green, G. H. — *Psychoanalysis in the Class Room*, University of London Press, Ltd., London, 1921

Hollingworth, H. L. — *Judging Human Character*, Appleton, 1922

Jastrow, J. — *Character and Temperament*, Appleton, 1923

Jung, C. G. — *Psychological Types*, Harcourt Brace, 1923

Koffka, K. — *The Growth of the Mind*, Harcourt Brace, 1924

Köhler, W. — *The Mentality of Apes*, Harcourt Brace, 1925

Saxby, I. B. — *The Education of Behavior*, Putnam, 1925

Stern, W. — *Psychology of Early Childhood*, Henry Holt, 1924

Thomson, G. H. — *Instinct, Intelligence, and Character*, Longmans Green, 1925

Thurstone, L. L. — *The Nature of Intelligence*, Harcourt Brace, 1924

Watson, J. B. — *Behaviorism*, W. W. Norton, New York, 1925

—— "What the Nursery Has to Say about Instincts," *Pedagogical Seminary*, June, 1925

Wells, F. L. — *Mental Adjustments*, Appleton, 1917

CHAPTER III

INTELLIGENCE: WHAT IT IS AND HOW IT IS MEASURED

A. What Intelligence Is

The Gestaltists view intelligence as the capacity to compact patterns of varying degrees of complexity. The patterns or configurations are discernible in both sensory and ideational fields and emerge into consciousness as a result of insight applied to a total situation confronting the learner. The situation is one of problem solving wherein a goal mediated by relevant stimuli is consciously sought. Habits, skills, processes, and attitudes are based on configurative learning, and levels of intelligence may be established by means of test problems whose degree of difficulty progresses from situations in which the configuration is obvious to those in which the configurative alignment of stimuli is less evident or is obscured by " detours."

Since the success in configurative learning depends upon the discoverable suitability of means to an end, it follows that the learner must possess the insight to discover their suitability and that the situation must yield a working configuration. But not all learners are so endowed and not all adjustments are so rationally

organized in context. One must recognize, then, a level of intelligence operating by compounding stimuli and responses into a complex in which either insight or pattern or both are lacking. This is trial-and-error learning, or the forced association of movements or ideas that characterizes unpedagogic teaching in such subjects as penmanship and the fundamentals of arithmetic.

For the Gestaltists, then, intelligence is manifested at a level lower than that manifested in configurative learning, possessing, however, many traits of the latter, such as the capacity to select and associate firmly those elements in a situation which facilitate goal finding and the ability to reinstate the whole chain of such forged association by an act of memory. The earthworm that incorporates in its movements a right or left turn and the puzzled or indifferent pupil who learns the multiplication table through sheer repetition, impressed more indelibly by punishment and reward, exemplify this type of adjustment.

The complete range of nonconfigurative and configurative intelligence embraces four levels, the levels at which learning takes place by accretion, assimilation, gradation, and redefinition.

Accretive learning is simply trial-and-error learning. In its initial stages it is devoid of understanding. It is repetitive learning facilitated by the instinctive trend of manipulation and the momentum of stimulus-response behavior.

The three other types are configurative in their

unforced association of the objects in the field of stimulation. Learning by assimilation envisages the total situation composed of allied stimuli and goal as forming a logical sequence. By an act of insight or understanding, only those stimuli are selected of the sensory field which exhibit a high degree of relevancy to the end sought. Such relationship is established, not by the random movements of trial-and-error learning, but by a sudden orderly, errorless, and continuous movement along the curve of progress. Köhler's apes learned to discern the configuration which inhered in the food placed out of reach in the confining cage and the sticks and boxes, whose potentialities as " tools " they discovered by an enlightening flash of insight.

Learning by gradation is configurative learning of a higher grade. Its organizing principle is " direction toward " of one stimulus to another where the degree of qualitative or quantitative difference is one of an ascending or descending gradient. Köhler's hens learned to take food placed on colored papers differing in intensity of color by discovering the gradations in the shades. Subsequent changes in the shade of one of the papers employed in the critical test served to influence the choice of that shade which sustained the same relative gradation as that which characterized the papers used in the training experiment. The apes' adjustment was made by discovering that the sensory elements and the goal were aligned in a coherent curve; the hen solved the problem by discovering in the total

situation the presence of a scale. Apes detect finer gradations in the intensity of stimuli and human beings still more refined differences. In the experiments performed by Köhler on apes and on a three-year-old child, the gradations constituted a more subtle configuration.

In learning by redefinition, enriched experience with numerous types of configurations enables the child to reassociate the relevant elements of allied configurations into more elaborate ones by causing the emergence of totally new patterns of behavior; the emergent configurations owe their vitality to the assignment of new meanings to the elements of configurations already learned while in quest of goals and the interaction of supporting stimuli of a like nature. The most elementary manifestation of the process of redefinition arises in connection with the so-called illusions of the senses. Thus the drawing in perspective of a cube may present the ambiguity of a solid or a hollow cube with three of its sides removed. Likewise a few short strokes like these may be interpreted as (1) the meaningless dabbling of a person in a state of vacuity, (2) a gun with bayonet fixed, carried on the shoulder of an invisible soldier, (3) a stroke of lightning tearing its way through a wall, or (4), to suggest another martial scene, a down-thrust bayonet penetrating an obstruction. In-

deed, these configurations do not exhaust the possibilities of redefining the situation. In ideational learning, redefinition enables the learner to make highly complicated adjustments in the field of moral and social behavior by applying old disciplines to new conditions and by discovering new implications in old principles.

But Köhler utters the warning that " even behavior with insight which indicates intelligence must not receive an interpretation too highly intellectualistic." Koffka characterizes the process of configurative learning as " intelligent manifestations of a primitive order," which break down, as we have seen, in the adjustment of the " detour-board " variety, involving a high degree of indirection in reaching the goal and the exercise of judgment and deliberativeness in organizing the field of perception and in inhibiting instinctive movements. Yet the fact of the awareness of a plan in the field of perception which may lead to the attainment of a goal argues for a degree of intelligence of considerable adjustive significance.

An animal, however, can reason only about things. He employs the rudiments of a " picture logic." He can discern the relation between things; he can recognize, albeit vaguely, their sequential nature. But it is only things that arouse this power; his thinking is done on the plane of perceptual inference. He cannot reason about ideas. A dog can make inferences about a man or men, but not about mankind. Moreover,

this facility in playing with ideas is largely due to man's power of speech, for without language abstract and speculative thinking would not have been possible.

On the physiological side, intelligence presupposes the possession of certain neural qualities. To have emerged from the trial-and-error stage, where blind gropings led to accidental success, the nervous system had to store up countless connections to enable the animal to experiment selectively, economically, and effectively. The memory unburdened itself of former half-successes, partial successes, and near-successes. The neural system bore the indelible marks of former "trials," so that the animal intelligence could "read the record of achievements in allied fields." But it did not "write" its impressions in promiscuous places. Experiences constantly formed neural connections which a directing force organized into systems and "hierarchies," and it is this systematization of past experiences that enabled the nervous system to carry its precious burden of memories free of a state of wild confusion. New experience was never suffered to go unrecorded for want of a blank page on which to write its story. In fine, intelligence depends upon neural connections:

1. There must be a great number of neural connections. The mind that meets every new experience in the spirit of an utter stranger cannot rise in intelligence. Many animals retain so little intelligence that every repetition of a task is a new adventure.

2. The neural connections must be organized, and this organization must attain some permanence. The nervous system must group and catalogue past experiences so that new experiences may be readily absorbed into a sustaining group.

3. The neural connections must be sufficiently flexible to facilitate ready readjustments. Intelligence must always remain on the *qui vive* for new experiences, new adjustments; it must keep its psychic capital liquid in order to make profitable investments in the mart of life.

We have seen that intelligence in its evolutionary aspects progresses in fine gradations from the intricacies of instinctive behavior through associative thinking, initiated by trial and error, to perceptual reasoning, from concrete to abstract through the intermediation of logical inference; and that, in the human race, mind and instinct have collaborated toward a refinement of the hit-or-miss experimental adjustments by means of new mental accessions, use of language, and thinking in ideas and concepts. This advance in mental efficiency is paralleled by a vast improvement in nervous organization, enabling the individual to store his past experiences in a manner that makes their manipulation easy; this equipment is attained by developing a *storehouse* of useful memories whose component impressions are deep-seated, effectively associated, and endowed with a nascent quality that facilitates the making of new combinations and set mutations. Specifically, what

renders the monkey's behavior more intelligent than that of the ant is the development of mental mechanisms that break the trammels of instinct and conduce to the progressive conquest of the environment. This conquest is due to the elaboration of rational powers whose motive forces are keen intellectual curiosity, dependable memory, and associative reasoning; the ant cannot add to its elaborate instinctive equipment while the monkey improvises new modes of behavior beyond the inventive capacity of mere instinct.

Significance of Intelligence for Educators

This finely articulated, highly complex mechanism is very significant for human progress. Though it is but one of the major traits of personality, its attributes are sufficiently distinctive to warrant individual treatment. Sociology is interested in the rise and fall of nations in so far as progression and retrogression are attributable to the psychic factors of civilization. The sociologist is concerned with the ameliorating influences of fine minds and the retarding influences of defectives. Education busies itself with the problem of educability. The teacher and the supervisor are attempting a solution of this problem by ascertaining in what ways the curriculum and intelligence interact; they are prospecting for mental forces with a view to harnessing them to the educational machine. The psychologist takes for his province the resolution of intelligence into its component elements, envisaging

reasoning as a fundamental capacity and grouping observation, perception, memory, and imagination as accessory aptitudes. In fine, the sociologist views intelligence from the point of view of social efficiency; the educator views it as learning efficiency; and the psychologist studies intelligence as means, adequate or inadequate, of compassing the aim of both sociologist and educator.

Obviously, we cannot measure intelligence unless we know something about it. The mental tester must have a definition of intelligence. A composite statement would accord due weight to the sociological, psychological, and pedagogical views.

What is intelligence? Stern defines it as "a general capacity of an individual consciously to adjust his thinking to new requirements." He emphasizes purpose, thinking, and adjustments. Meumann phrases it differently, but is in substantial agreement with Stern in stressing thinking, qualifying it, however, as independent, original, and productive. The saliency of the thought process receives emphasis in Terman's view: "An individual is intelligent in proportion as he is able to carry on abstract thinking." Lester Ward precipitates the concept "intelligence" and finds that it is composed of two ingredients, intellect and knowledge.

Educators have evolved more satisfactory definitions of intelligence. Binet's contribution has the merit of practicality; it defers, unconsciously perhaps, to the

popular notion of intelligence, but does not depart from the psychological treatment of consciousness.

Intelligence enables the individual to seize the world external to him and labors to reconstruct it out of the fragments that come into his possession. These fragments constitute the bricks which intelligence fashions into a psychic edifice. Binet calls these " bricks " the element "a," with a second element, which he calls "b." For the construction of the figurative edifice, various operations intervene; comprehension, memory, imagination, judgment, and language advance the structure to its completion.

Their joint action welds the element " a " to element " b," the intellectual mortar and iron and lumber, through the operation of invention, direction, and criticism. Reverting to psychological terminology, the adult must comprehend a given situation, fill in the lacunæ with experiential material, and apply the directive force of adopting means to an end under the guidance of a critical faculty intent on promoting efficiency. His invention is poor. Show him a picture of an unfortunate couple seated on a bench (a Binet test), and he reacts unsympathetically and impersonally. He sees a " man seated on a bench and there is a woman near him." He does not proceed beyond cataloging the component parts of the situation; he has no eyes for the shadows cast by them. Nor can the child direct his thinking nor sustain his train of thought. He is continually changing direction, lured

by the beckoning phantasy or the seductiveness of the unexplored bypaths.

These three aspects of intelligence — direction, invention, and criticism — disclose different degrees of intensity corresponding, roughly, to three levels. In the course of his psychic development, the normal individual comprehends, directs, invents, and criticizes data furnished by experience on the (1) perceptual, (2) representative, and (3) symbolic planes. He sees life with his physical eye, with his imaginative eye, with his speculative eye. He envisages a fact as a disparate entity, as a complex of sharply defined qualities, and as a symbol, prisoning in depths a spiritual Golconda yielding its treasures to the one possessed of the magic key.

Ward stresses that which was implicit in the definitions of Stern and Binet. Knowledge and intellect are closely related. From childhood there is formed a close union between the two because knowledge is the grindstone upon which intelligence whets its cutting edges. Indeed, it is this storing up of organized, well-ordered facts that spells growth in intelligence. These facts, these memories of successful experiments, emerge when heredity and acquisition make their first contacts. Even as the first glimmer of intelligence illumines the infant's face, environment has already made its indelible impress upon the child's innate mentality and every hour registers new impressions. Innate intelligence, therefore, is for the educator and mental tester a mere abstrac-

tion. For what is innate is not intelligence, but the capacity to become intelligent. Yet although we do not know how to measure it directly, we may still appraise its worth indirectly by testing the mind's affinities for knowledge; here "knowledge" is employed in a widely comprehensive sense to include intelligent behavior manifesting itself in knowing, in learning, and in coördinating accessory bodily movements.

Colvin frankly concedes this indirection in testing intelligence. He defines intelligence "as a group of innate capacities by virtue of which the individual is capable of learning in a greater or less degree in terms of the amount of these innate capacities with which he is endowed"; and is in full accord with Buckingham in holding that "we are justified, from an educational point of view, in regarding intelligence as ability to learn, and as measured to the extent to which learning has taken place or may take place." We must then regard intelligence as a fusion of innate and acquired aptitudes operating through the processes of learning, energized, as we shall learn, by volitional and emotional traits and measured by the ease, permanence, variety, and complexity of the acts, habits, and methods already acquired.

B. How Intelligence Is Tested

With a workable definition as a starting point, psychologists have devised tests in the form of tasks arranged in the order of difficulty and designed to repre-

sent (1) a learning situation, (2) a problematic situation demanding reasoning for its solution, and (3) a practical situation variants of which are encountered daily in the normal life of children. At first psychologists were confronted by an embarrassment of riches: there seemed to be no end to the number and complexity of suitable tasks. But gradually there emerged selective principles which relate to (1) universality of appeal, (2) independence of school learning, (3) independence of special talents, (4) degree of susceptibility to coaching, and (5) number of mental functions involved. To give a test involving a high degree of drawing ability would discriminate unduly against those untalented. To duplicate scholastic tasks would favor the child with educational opportunities. To stress reasoning would militate against the child with weak reasoning powers. Although the danger of coaching has not been entirely removed, the practice of employing duplicate forms has reduced it to a minimum.

(1) *Kinds of Tests*

Tests may be classified as individual and group tests. The individual test is also known as a scale, because it follows the principles underlying the construction of scales. Its name indicates that the examination is individual; only one child is examined at a time. The tester is known variously as examiner, psychologist, or mental tester. The child tested is known as subject, reagent, or examinee. The most popular forms of the

intelligence scales are the revisions and extensions of the Binet-Simon Scales of 1908 and 1911.

The group tests are used in testing groups of children. " Mass tests " is another name for them. These are subdivided into (1) composite tests, (2) synthetic tests, and (3) reasoning tests. The composite tests are composed of a number of subtests. A palpable attempt is made to retain the typical tests found in the original Binet Scale. Many new types have lately been introduced. For instance, many group tests have (1) directions tests, involving the execution of several commands, (2) arithmetical reasoning and arithmetical manipulation tests, (3) tests requiring a child to define concrete or abstract words, (4) word-opposite tests, (5) analogies tests, as, " The glove is to the hand as the shoe is to what? Glove : hand : shoe : ? " (6) memory test for permanency and extent of neural connections, and (7) range of information test to appraise the quality and quantity of nascent interest in the world external to the child. The purpose that animates the devisers of these tests is to provide enough variety to prevent duplication so that all mental powers are forced to contribute a share to the total score. Most of the American tests belong to this category.

The synthetic tests are based on a frank recognition of the psychology of mental functions. While their authors scrupulously reject the compartmental hypothesis of " faculty " psychology, they do not reject the schematic aid of its phraseology. They consider intel-

ligence, in its educational and vocational aspects, as effective associations of such mental capacities as memory, observation, association, imagination, concentration, judgment, and reason. Unlike the other group tests, the identity of the various subtests is not lost in the composite score, and this practice furnishes a percentual evaluation of each mental function.

Moreover, the apportionment of intelligence lends itself to a finer gradation of individual mentality. Comparison between the various essential phases of the mind's activities of the subjects may find a more significant expression than the statement of the difference between the respective scores. For the scores merge their identity, and the examiner is left with an interrogation as to the relative intensity of the mental qualities which operated to produce the difference. But the aforementioned tests afford the opportunity of equating differences with mental instead of statistical values. Dr. Schlag cites the cases of H and G. Their standing is summarized thus:

	Maximum	H	%	G	%
Attentiveness	75.5	75.5	100	54.1	71.4
Memory	46.5	32.5	70.3	16.5	35.5
Free association (Combination ability)	132	106	80.3	131	99.2
Comprehension	51	46	90.2	41	80.4
Judgment	100	69	69	100	100
Observation	73.5	58.5	79.6	53	72.1

H and G score 387.5 and 395.6 respectively. The difference is not significant, yet an analysis of their mental profile shows little mental correspondence. H is decidedly weak in judgment and inferior in linguistic ability (combination ability), while G has a poor memory but a fine power of judgment. The value of such detailed characterization for diagnostic and prognostic purposes is evident.

The reasoning test is a modified form of the synthetic tests. All mental functions are placed in a psychical hierarchy with reasoning at the apex. It is argued that the examinee who can reason through a problematic situation quickly and cogently can do so by virtue of keen observation. English psychologists are using many variations of this type of test.

Yet another classification is sometimes made. The intelligence test is differentiated from the psychological test. While the former surveys the entire mind, the latter confines itself to the quantitative evaluation of primary forces, such as sensory discrimination, immediate recall and delayed recall, and general memory and special memories. It affords valuable aid in diagnosing disabilities in school subjects due to the pupil's inferior sensory equipment or defective memory. A poor memory for numbers may explain the indifferent mathematician.

The nature of the responses evoked by the tests has an important bearing upon their reliability. Intelligence tests therefore are valued in proportion to the demands

made upon the linguistic equipment of the examinee; and tests are divided into verbal and nonverbal, linguistic and nonlinguistic. Many of the nonverbal tests are called performance tests, because the examinee is required to execute some movement or to arrange some objects. Such tests favor the linguistically inept or the physically handicapped, for a poor vocabulary is a nonconductor of the current of intelligence and hesitancy effectually shuts off its flow.

(2) *The Construction of a Scale or Test of Intelligence*

The terms " test " and " scale " are not interchangeable terms. The test measures vertically while the scale measures horizontally; the one measures by approximations to the maximum score; the other, by the extent of the approach to definite units called age-groups, the approach being limited, within wide boundaries, by the chronological age of the subjects. Individual tests of intelligence are scales; group tests are usually called tests. " Scale " connotes levels; " test " connotes maximum performance. The scale is applied at the point of greatest correspondence with the chronological age; test disregards age, it may be given to a wide range of ages, usually at the optimal age of school progress, progress being measured in terms of linguistic achievement. The lowest age is the age at which the subject can interpret the printed page. Many tests, however, have separate forms for the primary and grammar grades.

Yet another distinction may be pointed out. When used in conjunction with the term " intelligence," the term " test " signifies a battery of tests, sometimes called subtests. In all other cases " tests " are either subtests or component units of scales, as test 1 of the six-year-old age-group. These distinctions will be recognized throughout this book, corresponding with the designations employed by the authors of the various scales and tests.

(3) *Source of Material for the Construction of Scales and Tests*

Both tests and scales draw unreservedly from Binet's Scales of 1908 and 1911. Even the Army Tests are under heavy obligation to Binet, though Binet's contributions have been more on the side of inspiration than of adaptation. His title, the " Father of Intelligence Testing," remains unchallenged. This does not minimize the yeoman work of psychologists like Cattell and Ebbinghaus or the corrective influence of a host of educational statisticians and revisionists who have placed mental testing squarely on the road to scientific evolution of mentality.

(4) *Material for a Scale*

The material of an individual scale consists of four kinds of tests:

1. Tests comprising tasks which all children have had sufficient opportunity to learn provided they pos-

sess the requisite capacity for learning. Some of these are (*a*) telling the names of dates, weeks, months, and years, (*b*) knowing one's own name, and (*c*) learning the name and location of the salient sense organs.

2. Tasks the performance of which does not depend merely on practice. Such queries as "What is the difference between a butterfly and a fly?" "What is the thing to do if it is raining when you start to school?" or "What's left out in this picture?" are representative of this class.

3. Tasks that depend on school learning. Among these may be enumerated such tests as (*a*) vocabulary tests, (*b*) arithmetical problems, (*c*) definition of abstract words, (*d*) naming sixty words in three minutes, (*e*) similarities and opposites tests, and (*f*) rearranging sentences in which the words have been disarranged.

4. Tests of ingenuity and reasoning (the Healy-Fernald Puzzle is one), as (*a*) making logical inferences or (*b*) solving such problematic situations as "A mother sent her boy to the river and told him to bring back exactly 7 pints of water. She gave him a 3-pint vessel and a 5-pint vessel. Show me how the boy can measure out exactly 7 pints of water, using nothing but these two vessels and not guessing at the amount."

(5) *Qualities of a Good Scale*

Repeated investigations have produced a body of principles and practices that no maker of tests can afford to ignore. Not all of them are of universal appli-

cation. Nor are they accepted without cavil. But they possess sufficient validity to merit respectful attention. Continuous revision is resolving the moot questions. The qualities may be characterized as statistical, pedagogical, and administrative.

1. Statistical Qualities of a Good Scale. When the form of the tests has been decided on, the investigator must comply with certain statistical principles. Statisticians have been made aware of the indefiniteness of the two statistical assumptions that the distribution of mentality follows the normal curve of probability, and that the majority of children are of average mentality. The indiscriminate testing of a large number of children does not give much promise. This procedure ignores three important conditions:

1. Experience indicates that the number of retarded children is greater than the number of advanced children.

2. There is no way of controlling the distribution of the ages of those tested. There may be few nine-year-old participants, a great many ten-year-olds, and a negligible number of twelve-year-olds.

3. The presence of highly gifted precocities and decidedly subnormal retardates may give very unreliable results, just as the presence of very low or very high scores may produce an unrepresentative average.

These difficulties have been obviated by giving the tests to a selected group and by giving each age-group of tests to children of three consecutive ages.

The investigator selects a group of substantial homogeneity. This is called the control group. This group may be found in schools where careful grading is the practice. Such groups may be reliably considered as pedagogically normal; and when care is exercised in the choice of the group so as to eliminate social and economic influences, a fair selection may be made. When counterbalancing factors have been duly recognized, when teachers' estimates are in agreement as to the uniformity and normality of pedagogic achievement, the investigator proceeds with the test.

The homogeneity of the most carefully selected control group is only relative, owing to the great variability in the rate of mental growth. The variability consists both in speed and extent. Children mature at different rates and in different mental directions; the extent covers a wide range. This unevenness is indicated by the mental age. But this mental age may result from passing all tests of consecutive age-group tests or a number of subtests in widely scattered age-groups.

This fact of scattered scores must be taken into account in determining the percentage of passes for each age, for the investigator must decide in which age-group the subtests belong. Kuhlmann cites two supposititious cases to illustrate the procedure. He assumes a scale with five tests for each group. He assumes further that he is going to test 100 selected children averaging six years, and that these children display the tendency

to pass tests in the seventh and eighth age-groups. There would result the following distribution:

Age-Group	Total Number of Passes (100×5 [Number of Tests] = 500)
I	500
II	500
III	500
IV	500
V	400
VI	300
VII	200
VIII	100
IX	0

If all these children had passed all the tests of the sixth-year age-group, they would have made 3000 passes, for they would have made 500 passes in years I, II, III, and IV, 400 in year V, 300 in year VI, 200 in year VII, 100 in year VIII, and no passes in year IX, a grand total of 3000 passes. Since this represents the number of successful passes of 100 six-year-old children of normal endowments, the test of this age-group properly belongs to the sixth year. Phrasing it differently, six-year-olds of this selected group average a mental age of six years.

The subtests are shifted about till their correct place is ascertained. When this placement has attained the stage of stabilization, the investigator proceeds to the stage of verification. Other selected groups are examined, and when a very close correlation is achieved,

norms are established. Thus the norm for an average six-year-old age is a mental age of six years, figured on the number of maximum passes of the corresponding age-group or scattered portions of succeeding age-groups.

2. Pedagogical Qualities of a Good Scale. The tests are tasks placed before children. They must conform to the pedagogic principles underlying the activity of the mind in its aspects of attention to stimuli comprehending its signification and of retention for a space, for the purpose of comparing and judging. To be pedagogically sound, a scale must possess objectivity and regard for the span of attention.

(*a*) *Objectivity.* There must be no doubt in the mind of both subject and examiner as to the meaning of the directions. The language must be of such unequivocation and transparency as to permit the child to come directly at grips with the problem. Anything that detracts from perspicuity of expression deflects the mind from the task and reacts upon the nature of the response. The rule is to speak or write in so simple and direct a form that no child who speaks English can have cause for misunderstanding. Unusual or abstract words must be eschewed; meticulous care must be exercised to exclude expressions that may suggest a humorous meaning or a slangy construction.

(*b*) *Regard for the Span of Attention.* The ability to follow directions involves a high degree of mental ability. This ability varies with the degree of maturity.

Therefore the test must not contain many units. Too many elements strain the memory span. The strain of holding captive many directions may well lead to loss of poise and to confusional thinking; and a child incapable of prolonged concentration either becomes apathetic or exerts its waning powers to the retention of chance portions of the instructions. Many investigators evade this difficulty by providing for repetition or trials, by splitting up the directions at logical divisions, or by allowing for extended explanation, though without crediting or penalizing the subject.

It may be well to examine a test that embodies some of the characteristics of a pedagogically sound test. Perhaps the test is one of giving word opposites. The directions are: "I am going to give you a word and see how quickly you can think of another word that means just the opposite. For instance, if I say 'full,' you must say 'empty'; if I say 'poor,' you must say 'rich'; if I say 'black,' you must say 'white'; and so on, as fast as you can go. Now what do you say when I say 'full'?" Repeat for "poor" and "black," giving the pupil the correct word again if he fails. Then use the list of ten words, giving him the correct word for the first three if he fails.

The test is worded in simple language. There isn't a superfluous word. It conforms admirably to the exigencies imposed by the limitations of mental concentration. It does not make too large a demand upon the child's power of comprehension. It affords ample

opportunity for leisurely apprehension, and it makes ample provision for circumventing possible misapprehensions by cross-examination. Further assistance is given through the first three steps of the test proper.

3. Administrative Qualities of a Good Scale. As an administrative instrument, the scale must possess a degree of accuracy compatible with its use in diagnosis, and an unquestioned ease in manipulation. The scale can be used as an educational thermometer only when it measures temperature with a fair degree of exactitude and when it can be read easily. The precision of a scale may be termed its discriminative capacity. Statistical validation merely furnishes broad sanctions. A scale must be refined by standardizing procedure, method of scoring, and method of recording responses.

(*a*) *Standardization of Procedure.* The problem of procedure has been partially solved by the constructor of the scale. This information may be found in the manual that accompanies it. In it the examiner is instructed how to give each test. But the instructions need supplementation.

1. Familiarize yourself with the tests. With the aid of the manual examine a number of children. Systematic examination is not necessary. Copy down all the answers. Abbreviate whenever possible. Invent forms of abbreviations and fix them in your mind through adequate repetition. Now lay aside the record of the child's responses and test again after a lapse of a day. Compare the results and account for discrepancies.

2. Learn how to create the proper atmosphere. Many children are discomfited by the examination. Practice will prescribe the method of approach when dealing with a diffident child, an overconfident child, and a stubborn child. The tests have been standardized with children under normal conditions of emotional stress. The examiner must strive to surround the child with the necessary safeguards against nervous strain, self-consciousness, and "examination fear." In the absence of these precautions the responses cannot be termed normal responses. In the event of the persistence of the unnatural attitude the examination should be postponed.

3. Maintain an atmosphere of success. Encourage the child by the bestowal of judicious praise and the abstention from expressions or intimations of censure. The younger the child, the more copious should be the praise. Extravagant praise is better than tepid commendation. "That's fine!" "That's very good!" help to put the subject at his ease.

4. Guard against fatigue and the flagging of interest. This can best be done by disarranging the order of the tests in each age-group so that (1) the easier tests come first; (2) tests requiring alertness or effort follow; (3) tests of routine nature, such as furnished by reading and counting, come next; and (4) tests involving motor activity come last. This procedure may be summed up in this rule: Work from the easier to the harder and from the verbal to the mechanical.

5. Seek out every means of adding to the score of the examinee. Give him the benefit of every chance credit that he might earn in other age-groups, so that he may get the benefit of a wide scattering of passes. It is advisable to begin the test at the nearest age-group that corresponds to his chronological age. If success crowns his first efforts, proceed to the point where no successes are registered. If the subject passes are widely scattered, it may be best to extend the range to the utmost limits.

Conversely, if the examinee does not secure any passes at age, work backward until a point is reached where all tests of the age-group are passed. Work forward to a point of undoubted continuous failure. Now assign to each pass its corresponding credit in mental months. Total the amount and enter it in the proper space.

6. When self-confidence has been attained, the examiner should adhere to a fixed procedure both in making records and in forging a subjective bond between himself and the examinee. In other words, there must be standardization of the examiner's procedure as well as standardization of the author's procedure, with flexibility of interpretation in the former and undeviating rigidity in the latter.

(*b*) *Standardization of the Method of Scoring.* The premier quality of a good method of scoring is that of objectivity. The test should evoke a short, simple, straightforward response that may be easily expressed

in simple language. The best response is one that may be evaluated quantitatively rather than qualitatively. Qualitative evaluation of responses introduces the disturbing factor of the examiner's judgment and unconscious bias. In the scoring of many tests the personal equation enters in varying degrees. Thus in the test in " pointing to parts of the body," the examiner is granted the latitude of giving credit " if correct part is indicated (in any way) three times out of four." In the test requiring the objects by " use or better," the examiner enjoys a still wider amplitude in interpretation, while in the test of defining in terms " superior to use " the examiner's judgment has far greater scope.

The assignment of credits should take the form of whole units. Partial credits are subversive of exactitude and may militate against the discriminative capacity of the test. " Wrong " or " right " is the only judgment easy to make. Fractionalization lends itself to breaches of neutrality in cases where the examinee's personality is prepossessive or obnoxious; it is not a very easy task to rid oneself entirely of personal predilections or antipathies.

The best method of scoring is that employed by Kuhlmann in his " time and error " tests. In counting dots on squared paper, the examiner merely records " the number of squares counted wrongly (errors) and the total time for counting all squares (time). Passed if $\frac{1}{10} + 5\ \mathrm{E} = 15$ or less." The denominator of the first term refers to the number of squares and the number 5

of the second term is the equated compensation of the time element. The whole process is predominantly mechanical.

The allotment of credits deserves serious consideration; and its seriousness is enhanced in tests involving trials. Of these there are three kinds: (1) tests of one trial only, resulting in complete failure or complete success, (2) tests giving several trials on the same task, and (3) tests giving several trials not of the same difficulty, counting as successes if one trial results in a pass.

This practice is of dubious validity. The contention that success in one trial proves ability, while failure in one or more is an indication of mental disturbance rather than mental incapacity, is answered by the equally plausible affirmation that the resistance to disturbances is a mark of intelligence and that the boy who passes at the first trial is of a superior stamp. However, there is but one sovereign method of arriving at a decision as to the adequacy for any scoring method: the investigator must ascertain empirically which method gives the test the greatest discriminative power. The question is not "Which is the best scoring method?" but "Which method discriminates most accurately the individual from the group?"

(*c*) *Standardization of the Method of Recording Responses.* This phase of testing has not received the study it deserves. Here the greatest divergence is to be found. A good method of recording responses should possess the following characteristics:

(1) *Simplicity.* The simpler the technique, the superior the method. Overelaboration but adds to the onerousness of the task and involves a consumption of time that but lengthens unduly the examination period.

(2) *Space for recording responses.* The quality of response being an important factor in diagnosis, ample space must be provided for noting the substance of the answers, to form a basis of comparison for future examinations.

(3) *Separate blanks for each age-group.* This makes for ease of scoring in scales where each test is accorded different values. Thus such simplification would constitute an improvement over the Terman Record Booklet, which contains tests of two months' credit and those of four months' credit. Where one blank is used the examiner is prone to overlook this difference, to the invalidation of the examinee's mental status.

(4) *Separate blanks for abbreviated tests.* Terman provides for brief examination forms. These are composed of the crucial tests of each group, for which added credits are indicated. Although these salient tests are starred, their inclusion in the body of the complete tests is often very confusing.

(5) *Blanks for supplementary tests.* At times the need of verifying the findings of one examination makes itself acutely felt. When the supplementary test is a group, the test space may be found on the examination blank under the heading "Other Tests." The subjoined form should give satisfaction. Only the head-

ings are printed. The rest is filled in by the examiner. The form is very valuable for exhaustive examination.

Name of Child....Age....Years....Months....Birth....School....
Grade.....Date.....Examiner.....Name of Test.....I.Q.....

SUPPLEMENTARY TESTS AGE 3

No. of Test	Scale	Test	Response	Credit
1	Terman	Pointing to parts of the body	H. Hesitant. Lacks directness I. Inattentive	
2		Naming familiar objects		
3		Enumeration of pictures		
4		Giving sex		
5		Giving last name		
6		Repeating sentences		
1	Kuhlmann	Enumeration of objects in pictures		
2		Pointing out parts of body		
3		Giving the family name		
4		Repetition of a sentence of six syllables		
5		Naming familiar objects		
6		Naming pictures from memory		
7		Tracing a square		
8				

(6) *Manipular convenience.* Manipular convenience is secured when the test is constructed conformably to the requirements just described. Facility of manipulation is further augmented by the compactness and arrangement of the test material. Blanks, pictures, diagrams, weights, and blocks should prove easy of accessibility. Each should be plainly marked with

legible numerals, indicating the age-group for which it is designed. Other things being equal, that scale should receive preference which has the less elaborate test material.

C. Educational Measures of Intelligence

(1) *The Educational Age (E.A.) and the Educational Quotient (E.Q.)*

The E.A. measures levels of educational growth. The instrument is the educational test. Age-norms are constructed for the major subjects of the curriculum. For example, there are norms in silent reading for various school years. The attainment of any pupil may be determined by comparing his score with the standard score for his age.

McCall has worked a composite E.Q. on the basis of seven tests: (1) reading, (2) completion, (3) addition and subtraction, (4) composition, (5) multiplication and division, (6) vocabulary, and (7) spelling. The scores in each subject were subjected to statistic evaluation, and each subject was weighted according to its contribution to the total of all scores. Each subject was then given a coefficient by which the score made in each subject was to be multiplied. The scores for each pupil were weighted and the resulting scores added to form the composite score. McCall has worked out norms for each grade. These norms are known as norm composites. Since the tests were made in May, he estimated the average ages for each grade for that

month. These average ages are known as educational ages, for the corresponding score indicates the educational status of that age.

The E.A. for every pupil is computed from a table of norms. For example, pupil A has a composite score of 91. According to the table the nearest norms are 71 and 107. Score 71 corresponds to an educational age of 115 months. Score 107 corresponds to an educational age of 128 months. This pupil's educational age lies between 115 and 128 months. His actual educational age is found by interpolation (McCall, *How to Measure in Education*, p. 37).

$$115+\left[\frac{128-115}{107-71}\times(91-71)\right]=115+(\tfrac{13}{36}\times 20)=122$$

The educational quotient, or E.Q., of pupil A is found by dividing his educational age by his chronological age. His chronological age is, say, 109 months. Therefore E.Q. is $\frac{122}{109}$ or 112.

(2) *Accomplishment Quotient (A.Q.)*

McCall has also proposed the A.Q. as a measure of actual achievement. It is found by dividing the pupil's subject-matter age by the mental age. There is reading age, arithmetic age, composition age, etc. For each subject, the average accomplishment for each age supplies the requisite data by which to judge the rate of a pupil's progress. "The A.Q," according to McCall, "is the most exact present-day measure of the efficiency of study, instruction, and supervision; it is

the only just basis for reporting to parents and for judging pupils."

(3) *Mental Age and I.Q.*

Statistics and psychology have formed a close union. It is sometimes difficult to tell where statistics ends and psychology begins. Both are measures. Statistics employs mathematical exactitude, while psychology draws its analogies from mathematics but does not base itself on mathematical certitude: it applies the reasoning of mathematics to the speculations of psychology.

The mental age is not a mathematical formula. It is an expression of equivalence in the form of a statistical figure. The mental age corresponds to a level of intelligence beginning in infancy and ending in maturity. But mental maturity is not coextensive with biological maturity nor with chronological age. The degree of mental maturity is determined from the scores obtained from mental tests and scales. Its relation with chronological age is that of a standard: the child of proven average intelligence passes the tests assigned to that age-group. The six-year-old passes all tests in the six-year age-group. Thus the various age-group tests indicate the degree of mental maturity and the intelligence quotient determines the degree of correlation between the chronological age and the mental age.

Computation of Mental Age. Not all children of a certain chronological age obtain a like mental age. Here is a marked disparity. A six-year-old boy may succeed only in passing the tests assigned to the fifth-

year group. The mental tester pronounces him a child with a mental age of five years. Should he pass some of the tests of a higher year, he may conceivably receive credit for a fractional part of an additional year, thus augmenting his mental age.

This augmentation cannot exceed sixteen years, for psychologists are in fair agreement that mental years correspond roughly to the chronological age of sixteen. This delimitation assumes sixteen levels of intelligence, which in turn are subdivided by means of the *intelligence quotient.* The intelligence quotient (I.Q.) is the ratio, expressed decimally, between the mental age and the chronological age. It is computed by dividing the mental age by the chronological age. Thus a child with a mental age of six and a chronological age of six has an I.Q. of 100; and one with a mental age of five years, six months, and a chronological age of five years, ten months, has an I.Q. of 94. This figure is obtained by reducing both ages to months and by disregarding the decimal point: 66 months divided by 70 months.

The I.Q. ranges between a theoretical zero and a maximum above unity and serves as an index of reciprocal factors of mental age and chronological age. There is a noticeable lack of agreement among psychologists as to the relative merits of mental age and the I.Q. Some would disregard the chronological age and use the mental age in classifying pupils. Dr. Van Denburgh puts his faith in the mental age. Dr. Kuhlmann prefers the I.Q. He argues that mental growth is

unequal at different age levels: the younger the child, the more rapid its mental development; the older the child, the less substantial the mental accretions. Mental age, then, does not mean the same for all chronological ages. The I.Q., however, remains fairly constant because the chronological age acts as a compensatory factor that neutralizes the unevenness of the mental age.

While some doubt the adequacy of either measure for purposes of administration, the I.Q. subserves the very useful purpose of introducing a scientific classification of mental levels. Its scientific accuracy depends upon the extent to which the intelligence test possesses discriminative capacity and the mental age it establishes measures a definite stage of mental growth. When this indispensable condition is fulfilled, the I.Q. indicates tangible levels of mentality beginning with idiocy and consummating in precocity. The entire range includes the following levels:

Level	Range in I.Q.
Idiots	0– 24
Imbeciles	25– 49
Morons	50– 74
Borderline	75– 84
Dull	85– 94
Average	95–104
Bright	105–114
Very bright	115–124
Superior	125–149
Very superior	150–174
Precocious	175 and over

Exercises for Self-Testing and Investigation *

I. Only (*a*), (*b*), or (*c*) will correctly complete each of the following statements. Indicate in each instance whether it is (*a*), (*b*), or (*c*).

1. Binet undertook to test
   (*a*) power of concentration.
   (*b*) the general level of intelligence.
   (*c*) memory and imagination.

2. Binet's psychology is
   (*a*) "faculty."
   (*b*) dynamic.
   (*c*) structural.

3. The most popular test of intelligence is
   (*a*) the National Intelligence Test.
   (*b*) Myer's Mental Measure.
   (*c*) the Stanford Revision of the Binet Scale.

4. The best way to characterize the intelligence of the child is by the use of
   (*a*) the descriptive terms "bright," "moderately bright," "dull," "very dull."
   (*b*) the descriptive terms "bright," "normal," "dull."
   (*c*) the I.Q.

5. Spearman conceives of intelligence as
   (*a*) a combination activity.
   (*b*) a common central factor.
   (*c*) a general capacity for adjustment.

6. In the matter of placement of tests, Binet discarded any given test if
   (*a*) 90% of a given age-group did not pass it.
   (*b*) 50% of a given age-group did not pass it.
   (*c*) the percentage of passes for a given test increased but little from the younger to the older.

* Planned for use with Chapter IV also.

7. The phrase "at age" is used in connection with
   (*a*) Burt's Revision.
   (*b*) Kuhlmann's Revision.
   (*c*) the Stanford Revision.
8. The most dependable type of intelligence test is
   (*a*) the individual test.
   (*b*) the group test.
   (*c*) the combination (intelligence and achievement) test.
9. "What's the thing to do when it is raining?" is a test of
   (*a*) memory.
   (*b*) imagination.
   (*c*) comprehension.
10. "Alternative tests" are employed in
    (*a*) Burt's Reasoning Test.
    (*b*) Porteus' Maze.
    (*c*) the Stanford Revision.

II. Binet stated that certain of his tests had the quality of stimulating the subject:

1. to "keep in mind the end to be attained, that is to say, the figure to be formed;
2. to try different combinations under the influence of this directing idea, which guides the efforts of the subject even though he may not be conscious of the fact;
3. to judge the combination which has been made, to compare it with the model, and to decide whether it is the correct one."

Arrange the tests in the Binet Scale in the order in which they approximate the above conditions.

III. One qualifying word or phrase is more pertinent to the general theme than the others. Indicate whether it is (*a*), (*b*), or (*c*).

1. Binet tests tend to bring to light
   (*a*) the child's special talents.
   (*b*) the quality of his memory.
   (*c*) the general level of intelligence.

2. Any accepted test of intelligence may be used as a
   (*a*) pedagogical guide.
   (*b*) sociological guide.
   (*c*) psychological guide.
3. The best method of grading intelligence is by
   (*a*) the use of standardized tests plus teachers' judgments.
   (*b*) the use of standardized tests plus pupil's history.
   (*c*) the use of standardized tests plus pupil's play habits.
4. Teachers usually
   (*a*) overestimate the intelligence of the older, retarded pupils.
   (*b*) are in substantial agreement with the findings of the intelligence tests.
   (*c*) assess equitably the intelligence of the older, retarded pupils.
5. An intelligence scale should contain a great many tests because
   (*a*) some tests may be used as alternatives.
   (*b*) there is a better chance of testing a pupil's power of concentration.
   (*c*) intelligence has many aspects.
6. Definite feeble-mindedness is indicated by an I.Q. of
   (*a*) 65.
   (*b*) 40.
   (*c*) 30.
7. A certain boy has a chronological age of 8 years, 10 months, a mental age of 11 years, 1 month, and an I.Q. of 130. This child is
   (*a*) normal.
   (*b*) bright.
   (*c*) gifted.
8. A feeble-minded child is characterized by the incapacity to
   (*a*) stand erect.
   (*b*) memorize large numbers.
   (*c*) give sustained attention.

9. A psychological examination should be conducted
   (*a*) in the classroom in the midst of other children.
   (*b*) in the classroom after children are dismissed.
   (*c*) in a specially equipped room and in seclusion.

10. The sequence of tests best calculated to place the pupil at his ease is
   (*a*) language tests and memory tests first.
   (*b*) those which test knowledge about objective things, like pictures, weights, and coins, and those which test judgments upon them.
   (*c*) naming sixty words, finding rhymes, making sentences.

Bibliography

Achille–Delman, F., and Ball, M. — *La Personnalité Humaine*, E. Flammarion, Paris, 1922

Carey, A. — "Factors in the Mental Processes of School Children," *British Journal of Psychology*, Vol. 8, p. 80, 1915

Conklin, E. G. — *Heredity and Environment in the Development of Man*, Princeton University Press, 1915

Fox, C. — *Educational Psychology*, Harcourt Brace, 1925

Goddard, H. H. — *Human Efficiency and Levels of Intelligence*, Princeton University Press, 1920

Gregory, C. A. — *Fundamentals of Educational Measurement*, Appleton, 1922

Ogden, R. M. — *Psychology and Education*, Harcourt Brace, 1926

Rogers, A. L. — "Intelligence Tests and Educational Progress," *Educational Review*, Vol. 61, pp. 191–216, February, 1921

Root, W. T. — "The Intelligence Quotient from Two Viewpoints," *Journal of Applied Psychology*, Vol. 6, pp. 267–275, September, 1922

Spearman, C. — "General Intelligence," *American Journal of Psychology*, Vol. 15, pp. 222–225

—— "The Heredity of Ability," *Eugenics Review*, pp. 219–237, 1914

Stern, W. — *The Psychological Methods of Testing Intelligence*, Warwick and York, 1914

Terman, L. M. — *The Measurement of Intelligence*, Houghton Mifflin, 1916

Thorndike, E. L. — *Educational Psychology*, Vol. 3, Teachers College, Columbia University, 1914

Thurstone, L. L. — *The Nature of Intelligence*, Harcourt Brace, 1924

—— "What Is Meant by Intelligence?" *Journal of Educational Psychology*, Vol. 12, pp. 201–207

Wallin, J. E. W. — "The Results of Retest by Means of the Binet Scale," *Journal of Educational Psychology*, Vol. 12, pp. 392–400, October, 1921

Woodrow, H. — *Brightness and Dullness in Children*, Lippincott, 1919

—— "The Nature and Measurement of Intelligence," *Journal of Educational Psychology*, Vol. 12, pp. 207–210, April, 1921

CHAPTER IV

REPRESENTATIVE INDIVIDUAL SCALES AND TESTS

A large number of scales are available for individual testing. For the most part they are revisions and modifications of the Binet scales. The exigencies of space do not permit exhaustive treatment. This chapter will concern itself with those of unquestioned merit. These tests may be roughly classified as principal scales and supplementary scales. The principal scales are the Stanford Revision (Terman), the Kuhlmann Revision, and the Burt Revision. The first two are American, and the last is English. The supplementary scales are the Porteus Maze, the Burt Reasoning Test (abbreviated form), and the Detroit Kindergarten Test.

A. Principal Scales of Intelligence

(1) *The Stanford Revision of the Binet-Simon Tests*

The Stanford Revision of the Binet-Simon Tests consists of a series of tests divided into age-groups. They cover ages three through eighteen. Each age-group up to the age of ten contains six subtests and one or two alternative tests. Each test counts as the equivalent of two mental months. Ages ten and eleven are combined and contain six subtests and three alternative

tests. Each is credited with four mental months. Ages twelve and thirteen are also combined, consisting of eight tests with mental credit of three months. Ages fourteen and fifteen are combined into a six-test age-group and one alternative, with mental credits of four months. A similar economy is effected in ages sixteen and seventeen; six tests and two alternatives represent this age-group. The mental credits are the same as for the previous age-groups. Age eighteen numbers six tests and no alternatives. The mental credits per test are two months.

1. Distinctive Features. This scale has an unprecedented vogue. Its popularity is due to the inclusion of many distinctive features:

1. The instructions are complete. *Measurements of Intelligence* is a complete guide to procedure, scoring, and interpretation. *The Condensed Guide* is a condensation of the former and is intended for the use of experienced examiners. *The Record Booklet* is an individual examination booklet, a pocket edition of the *Guide*. It has ample space for recording and scoring responses. An Abbreviated Filing Record Card furnishes a complete filing record for each child.

2. The test contains many excellent subtests of proven discriminative capacity. Of these the reasoning tests are the best.

3. The test material is simple and compact, and the pictures and drawings are printed on durable boards. The lines are deep, clear, and distinct.

4. It recognizes the possibilities of internal grading. Tests reappear in many age-groups, the magnitude of the score determining its age-assignment. This form of grading affords excellent opportunity for evaluating the growth of fundamental capacities, such as memory for numbers and vocabulary range.

2. Defects of the Scale. The scale suffers from some defects, the removal of which should enhance its discriminative capacity.

(*a*) *Marked Linguistic Nature.* Too many tests hinge upon the examinee's interpretation of the language in which it is couched; and many of the responses call for considerable linguistic ability. Indeed, paucity of vocabulary accounts for many failures and the consequent adverse I.Q. rating. What is not comprehended is not apprehended, and apprehension waits upon expression; nowhere is command of vocabulary more urgent than in the definition of abstract terms. Terman tacitly admits the saliency of the linguistic element when he observes that " in a large majority of cases the vocabulary test alone will give an intelligence quotient within *10 per cent* of that secured by the entire scale."

(*b*) *Lack of Uniformity in Procedure.* The examiner is tempted to culpable laxness by the author's too ready assent to liberal modifications of procedure. Crucial tests constitute an inviting abridgment, and the accommodation alternative tests, though admittedly of inferior quality, attract by their simplicity. Their free use detracts immeasurably from the precision of the

scale. It is as if one were to use interchangeably a yardstick, the span of one's hand, or the span of one's foot. The best of scales is of questionable precision. To allow modifications is to undermine its scientific foundation.

(*c*) *Need in Some Tests for Further Standards.* Many of the tests are not clearly expressed, and examiners are forced to elaborate the instructions. Take the ball and field problem. It is not made clear that the small circle represents a large field. The problem also calls for a decided revision of the examinee's notion of baseball fields. His experience demands a diamond-shaped or, at least, an oblong field. A circular field has almost the disconcerting force of an anomaly.

The wording is not very clear: " Let us suppose your baseball has been lost in this round field. [1] You have no idea what part of the field it is in. You don't know what direction it came from, [2] how it got there, [3] nor with what force it came. [4] All you know is that the ball is lost somewhere in the field. . . ." Statements [1], [2], [3], and [4] call for a large memory span, beyond that possessed by the average eight-year-old child. The strain of remembering these instructions is augmented by the unfamiliar aspect of " direction" and " force."

Nor are the solutions, as between inferior plans of the eight-year-old and the superior plan of the twelve-year-old, of such marked divergence as to merit the twofold age-assignment. The rough but continuous

spiral does not really surpass in ingenuity the allegedly superior plan of concentric but discontinuous circles; and many of the possible solutions indicated by the score card are merely fantastic and only rarely encountered.

(2) *The Kuhlmann Revision* (1922)

The Kuhlmann Revision is another revision and extension of the Binet-Simon Scale. The original 56 tests of the Binet Scale have been increased to 129, with only 37 of the original 56 retained. It has been standardized with meticulous care. Seven thousand children and adults were examined. The age-groups range from three months to fifteen years. There are excellent tests for age-group between three months and three years that require for responses some form of sensory acuity, motor coördination, imitation, and memory for common objects. For ages three to thirteen the author has devised age-groups of eight tests each. The tests show a marked diminution of the linguistic influence, but a predominance of motor responses; they are virtually tests in "following directions." Linguistic responses are conspicuous for the absence of extended or invaluable explanations. For ages thirteen to fifteen, eight tests are provided. These are graded in difficulty, and constitute, in effect, a measure of the fundamental capacities involved in general intelligence — a tacit reversion to "faculty" psychology. Thus the tests entitled "Simple Arithmetical Problems" and "Drawing Triangles on a

Square According to Directions" are substantially tests in observation and concentration. "Drawing Upright Forms in Inverted Positions" is a test of imagination; "Following Directions in Confusing Text" is a test of distracted attention and suggestibility; "Locating Sections of a Square from Description" is a test of association; "Immediate Recall of Unfamiliar Forms" is a test of visual memory; and "Making Logical Inferences" is a test in formal reasoning.

1. Distinctive Features. The Kuhlmann Revision possesses notable features which may be considered improvements on the Stanford Revision.

(*a*) *Marked Objectivity.* The responses fall within a definable range, predominantly nonverbal and easily scored. The "time and error" tests minimize the influence of the personal equation and achieve the detached and impersonal evaluation of the group test.

(*b*) *Greater Discriminative Capacity.* Eight tests are more searching than six. Statistically, it is demonstrable that the greater the number of units in a scale, the fairer the grading and classification. Grades A, B, and C permit wide variation in the assignment of borderline estimates; and personal judgment has the widest scope between excellence, mediocrity, and failure. The larger the number of units in the measuring instrument, the narrower are the limits for the play for subjective estimation.

(*c*) *Explicit Directions for Giving Tests.* While the examiner is not converted into a questioning and scoring

automaton, the directions anticipate special circumstances and indicate the right procedure in explicit, detailed instructions. The result is that the examiner is emboldened to face any case, and the examinee is not forced to divide his mental energies between filling in lacunæ in the directions and attacking the problem.

2. Defects of the Scale. The defects consist in the tentative nature of the tests for pre-school years and in the elaborate system of scoring. Many of the time and error tests make excessive demands on the examiner's time. Nor are they less exigent for the examinee. The directions require for their comprehension a large memory span and full concentration. In this regard these tests are really more than tests of intelligence; they are also tests of emotional stability and industry.

(3) *The Burt Revision* (1921)

The Burt Revision is an adaptation of the Binet-Simon tests for English children; and the peculiarly English adaptations are concerned with the substitution of English names and the conversion of French money into English values. There are only four of the latter, and their inclusion does not preclude the scale being used for American children, as substitutions may readily be made.

The Burt Revision retains the best features of the Binet Scale. The author has reëvaluated not only the tests, but also the very language of the original; and Dr. Simon, the only survivor, has lent his aid in resolving

moot points of translation and interpretation. Dr. Simon's suggestions in the matter of translation have been followed: "The best translation is the one that leaves the test in the same age-group where it had been placed." The tests of the Scale are assigned to age-groups from age three to age sixteen. The tests are also numbered consecutively to facilitate the scoring of sporadic passes. Altogether, there are 65 tests. There is a conspicuous lack of conformity in the assignment of the number of tests. The number varies, the highest being in the six-year group and the lowest in the twelve- and thirteen-year groups. There is a key, in tabular form, for converting test scores into mental ages. The mental equivalence is worked out in fractions of a mental year, from $\frac{1}{12}$ to $\frac{12}{12}$, with their corresponding decimal values.

1. Distinctive Features. Characteristically, Burt has attacked the problem with thorough scientific detachment. Practical success has not blinded him to theoretical shortcomings of his own Scale; and he has turned upon it the powerful searchlight of self-criticism during the process of construction as well as during the period of verification. It is a rare occurrence in authorship — the creative impulses under the merciless analysis of the critical faculty. His critical assaults were prompted less by empirical or theoretical considerations than by mathematical calculations of extreme refinement. There emerge from this study a number of disturbing conclusions that leave one with an acute inter-

rogation as to the soundness of Binet's conception of intelligence testing. These will be more fully discussed elsewhere.

The distinctive features of the Burt Revision are many. The most notable of these follow:

1. Exact references are made to allied procedure in other scales. The examiner may note the points of difference so that he may be guided in imperative modifications of the tests. The consciousness of subtle nuances in the formulation of questions and the evaluation of responses leaves one with a sense of the seriousness of the problem of testing; and one is brought into close intimacy with competing scales — a relationship that results in a sense of mastery over the material.

2. Superior facilities are provided for testing defectives. The scale was tried out on a large class of conceded normality and deficiency, and the tests were later arranged in the order of difficulty. The easiest tests for normals are tests involving critical perception. The easiest tests for defectives are tests of suggestibility and picture interpretation. Tests that offer obstacles to the defectives are (1) scholastic tests of a linguistic nature (dictation, reading), (2) tests of primary memory (recalling numbers and syllables), (3) reasoning tests (differences, definition, mathematical reasoning), (4) tests of facility in the manipulation of words (mixed sentences, rhymes, sixty words in three minutes), and (5) tests of information of a scholastic nature (naming months or days, counting backwards). That linguistic

ability is the *pons asinorum* of the defective is a fact of grave significance for school administration and possesses some importance for the examiner. In the absence of mitigating conditions the examiner is justified in considering marked linguistic disability a presumption of feeble-mindedness. It makes it incumbent upon him to give tests of many consecutive age-groups in order to divulge stigmata of feeble-mindedness. This caution is especially applicable to borderline cases.

3. Dr. Burt has formulated norms for defectives that are at variance with those prevailing in this country. Instead of employing the I.Q. as the demarcating line — a delimitation of dubious exactitude owing to its inconstancy in defectives — he postulates the discriminative capacity of the mental age. In deficiencies in intelligence not coupled with temperamental instability, a child, to be judged defective, must be backward by at least:

2 years at the age of 6 to 7
2½ years at the age of 8 to 9
3 years at the age of 9 to 10
3½ years at the age of 12
4 years at the age of 13 to 14

Less exact but more general is the conclusion that deficiency may be presumed when the amount of mental retardation is equal to three-tenths of the calendar years. Thus a ten-year-old is a suspect of feeble-mindedness if his mental age is 7 years.

2. Defects of the Scale. The defects apparent are those suggested by an a priori examination, and some of the strictures are directed against the pervading English temperament.

1. It is not yet available in handy forms. Both tests and accessory charts are but part of a large volume dealing with somewhat extraneous subjects. *Mental and Scholastic Tests* is too unwieldy a volume.

2. The distribution of tests is widely uneven. The paucity of tests for the upper levels of intelligence is too apparent to require comment.

3. There is an unconscious bias toward gauging mental deficiency. The author's signal success in the field of mental deficiency has given the Scale a wider orientation at the level of feeble-mindedness. This may account for its attenuation at the other extreme, supernormality.

(4) *Comparative Survey of the Three Scales*

It might be of interest to discuss briefly the sharp lines of cleavage between the scales. Only exhaustive experimentation can reveal the degree of discriminative capacity attained by each; and such investigation is urgently needed to evaluate the scientific soundness of mental tests. It is the province of the disinterested research worker to place mental testing on the same high plane that modern industrial efficiency methods have attained.

1. Adherence to the Binet Traditions. Scales differ in the degree to which the Binet tests have been incorpo-

rated. The original versions have been in the main adaptations to suit American culture. Burt follows Binet very closely and his tests are characterized by a fidelity to the language of the original that is lacking in the others. In this connection Dr. Simon's aid has been invaluable. Terman is less imitative and more revolutionary. The Terman Revision is frankly an extension; he has departed from the original with an easy conscience. The Kuhlmann Revision is the most radical of the three. He has discarded many of the original tests, and the introduction of time and error tests marks the point of greatest divergence.

2. The Allocation of Tests. The three scales show wide difference in the matter of assigning tests to age-groups. Comparison is not always possible because of the many new tests introduced. For example, Binet assigned the test requiring the child to point to different parts of the body to age-group 3, in which assignment Burt and Terman are in full accord, while Kuhlmann places it in age-group 6. "Comparing Faces" is assigned to age-group 4 by Binet, to age-group 6 by Kuhlmann, to age-group 5 by Terman, and to age-group 4 by Burt.

3. Differences in Procedure. There is a great difference in the wording, the number of trials allowed, the degree of explanation permitted, and the assignment of partial credits. Each Revision must be scrutinized for variations in procedure. The test called "Comparing Faces" may be cited in illustration.

	Instructions	Test Material
Terman		One card containing six faces arranged in three rows
Kuhlmann	"Which is the prettier of the two?"	Three cards of two faces
Burt		Same as Kuhlmann

	Instructions	Alternative Procedure	Test Material	Scoring
Terman	"Which of those two is the prettiest?"		One card containing six faces in a row	Passed if all 3 responses are correct
Burt	"Which is the prettier of the two faces?"	Alternative questions suggested with proviso that responses do not count	Three cards each having two faces	Passed if responses are correct in first attempt
Kuhlmann	"Which is prettier of the two?" "Which one do you like best?"	If only 1 error is made for 3 cards the first time, repeat same order	Three cards of two faces each	Passed if first 3 answers are correct or 5 out of 6

The differences indicated are more vital than cursory reading indicates. There are minds that may quibble at the implications of comparative and superlative forms of "pretty." The word "like" may be evocative of personal predilection. It is conceivable that a

child may prefer an ugly face because it excites his risibilities or because it harmonizes with his artistic sense. Again, alternative procedure may provoke responses that are not at all comparable. The wide sweep of associative impulses may inject into the reaction totally unexpected and incalculable elements; and while the responses may be outwardly incorrect, it might well be in complete accordance with the examinee's understanding of the question. Nor is there no significance in the various methods of displaying the cards. The separate card escapes the temptation to disperse the attention over the entire card, to the palpable weakening of concentration. Besides, the multiplicity of faces invites comparisons between widely separated faces; and doubt arises as to whether the adjective "prettiest" could possibly be applied to some of the faces.

4. Limitation of the Binet Scales. The Binet conception of intelligence and its manifestations has dominated all the revisions. It assumes an innate intelligence and a superimposed acquired intelligence, merging imperceptibly into one fundamental capacity in which both lose their separate identities. The innate factor acts as a power while the acquired intelligence manifests itself as a force. Since "force" is characterized by concreteness and since it comprehends "power," it is advisable to consider the measurement of acquired intelligence as an indirect way of measuring innate intelligence.

General intelligence or mental maturity is a composite of three preponderant elements: innate ability, the source of mental maturity; education and environment, the maturing ferment; and age, the magnitude of this ferment. Each of these Binet proposes to weigh in formulating the mental age. But he assigns to innate intelligence the dominating rôle.

Burt has subjected the three factors to mathematical treatment, and he finds that a child's achievements in the Binet-Simon tests (including the revisions) is predicated largely by scholastic achievements. By applying the formula for partial correlations he finds that " of the gross result, one-ninth is attributable to age, one-third to intellectual development, and over one-half to school attainment. In determining the child's performance in Binet-Simon Scale, intelligence (innate) can bestow only little more than half the share of school and age but one-third the share of intelligence."

The implication is plain. The Binet tests must be supplemented by others dependent to a much lesser degree on school ability; tests must be found to compensate for failure in linguistic tests due to inadequate school training or training of a scholastic nature whether imparted by school or other agencies. Corroborative evidence is especially urgent in detecting serious deficiency and supernormality where glibness of tongue may mask the former while lack of that quality may hide the latter.

B. Supplementary Scales of Intelligence

(1) *The Porteus Maze Test*

The Porteus Maze Test accepts as a fundamental of testing intelligence that educability is the fairest index of intelligence. The ability to thread a maze is symptomatic of intelligence among human beings as well as among animals; and, it is argued, what proved successful with animals should prove adequate for testing children. The nonverbal nature of the Test is inferable from the fact that the scale of the testing material was originally designed for investigating educability in animals.

Dr. Porteus of Vineland, New Jersey, is the author. There are eleven mazes, each double tracked to simulate paths, which the child is invited to traverse from a marked point of ingress to a point of egress. The maze grows increasingly intricate as the test advances in difficulty. There are eleven mazes, for ages three to twelve and for age fourteen. They form an age-scale similar in significance to the Binet Scale age-groups. The child is confronted with a blank upon which a maze is drawn and is told to " suppose that this is a plan of the paths in a garden." These lines are walls which you cannot get over. Start with the pen " or pencil from the mark at the top (a capital S at the point of ingress), and find your way out of the garden by the quickest paths. Show me, first of all, any openings that you can see. (The two openings are indicated.)

All other paths are blocked. Don't go up any of the blocked turnings. Go down this path from the top, and then out by the *first* opening you can find."

The child is not permitted to retrace his steps to avoid blocked exits; he is directed to go back to the beginning. The trials are scored on a descending fractional scale. Thus, in orienting himself in the maze for year 14, the child is credited with 2 marks for a successful first attempt, $1\frac{1}{2}$, 1, and $\frac{1}{2}$ marks respectively for successes at the second, third, and fourth trials. For the years 3 to 4, inclusive, demonstration is permitted; the child is also allowed two trials and two or more errors. Where scores are not specifically indicated, one mark is assigned to each test correctly performed on the first trial and half a mark when a second trial is needed. Four marks are added to each total score to compensate for the presumable success in imaginary tests for ages 3 to 4, the test for age 3 being easy. The grand total gives a rough mental age.

The mode of attack is as significant as the quantitative result, for failure may indicate either intellectual immaturity or emotional disturbance. The child may be weak in systematic attention or he may be unable to plan ahead. Again, he may be too confident or too cautious: he either plunges *in medias res* or vacillates in the face of unanticipated moves. The tests are particularly efficacious in resolving doubtful cases of dullards, who fail hopelessly in the Binet tests and succeed with the Porteus Scale. Conversely, many

children possessed of plausible glibness fail in the maze tests; and their failure is due to their inability — amounting to serious deficiency — to be forehanded in planning moves and to maintain an attitude of alertness and attentiveness.

(2) *Reasoning Test* (*Short List*)

The Reasoning Test is an abbreviation of Burt's Graded Reasoning Tests. It is designed for older and bright children, and reveals, as far as one scale can reveal, the strength of logical thinking. It is inferential reasoning from given conditions based on facts of wide generality. Do they not include school learning or technical knowledge?

The series contains seventeen problems, arranged in order of difficulty and assigned in groups of two to successive mental ages. There are no age-groups below seven years; the upper limit is fourteen years. Children are asked to read each problem from a printed list or from individual cards on which one problem is pasted or typed. The instructions are: "Will you read this, please? At the end you will find a question. When you have read this question, look carefully again at what is printed above, and try whether you can think out the answer." The child must give both answer and reasons; therefore four successive trials are permitted. Each child works forward until he fails in three successive problems. A mark of 1 is given for each test correctly answered and correctly reasoned.

$\frac{1}{4}$, $\frac{1}{2}$, and $\frac{3}{4}$, respectively, are deducted for an ill-expressed reason, poor reason, or no reason. The children are then ranked according to the following table of norms:

Age Last Birthday	Average	Bright	Superior
7	2	5	8
8	4	7	10
9	6	9	12
10	8	11	14
11	10	13	16
12	12	15	17
13	14	16	17
14	16	17	17

The tests for age-group 7 are here reproduced:

"Kate is cleverer than May. May is cleverer than Jane. Who is the cleverest—Jane, Kate, or May?"

"I have bought the following Christmas presents: a pipe, a house, some music, a box of cigarettes, a bracelet, a toy engine, a bat, a book, a doll, a walking stick, and an umbrella.

"My brother is eighteen; he does not smoke, nor play cricket, nor play the piano.

"I want to give the walking stick to my father, and the umbrella to my mother. Which of the above shall I give my brother?"

(3) *Detroit Kindergarten Test**

The Detroit Kindergarten Test attempts to "differentiate the bright, average, and slow pupils" upon

* Published by the Public School Publishing Company, Bloomington, Ill.

their admission to the kindergarten. The authors, H. J. Baker and H. J. Kaufman, have designed an individual rather than a group test " because (1) the group test method is relatively less feasible for children so immature; (2) the measurement of intelligence of the young child is very important as predictive of classification in school; (3) the examiner has an opportunity to observe the nature of the reaction in addition to the score itself."

The Test is still in a formative state and is free from the rigidity characteristic of examinations which both standardization and statistization have stabilized. The authors temperately observe that it is " merely a starting point for further personal observation," and it is this very flexibility that enhances its value.

There are five tests, consisting of pictures arranged in rows; these vary in number from three to six. The child responds by pointing, and cross-examination — a valuable adjunct to any test — evokes short reënforcing responses. Test 1 is an information test. Test 2 requires the examinee to compare four rows of common fowls (four to each row) with the top row, which consists of one less specimen. " See these things here," the child is told as the examiner points to the sample top row. " Show me the one here (pointing to each of the succeeding rows) that is not there." The sense of logical relation is evaluated by means of Test 3. One row represents a broom with a broken handle, the remaining portion of the handle, and a handle ordinarily

inserted in a shovel. The child is invited to point to the " thing " that does not, logically related, belong with the other two. Test 4 is one of analogies. Three rows of objects are shown divided by a line into groups of two and three. One row consists of a bicycle, a car flanked by a dog, a wagon, and an apple. The child's reaction consists in perceiving the similarity between the first two and the wagon of the second group. Test 5 consists in the detection of similarities and differences. Test 6 is really a modified form of Test 3. Test 7 is an absurdities test. The child is faced with six " things " that deviate from the usual: a teakettle with two spouts, a boy displaying a clock where the head should be, a house with a misplaced chimney, a shoe with a double row of buttons, a child sleeping at the wrong end of a bed, and a hand with a supernumerary sixth finger, clutching a tooth brush. Test 8 is a reasoning test. It consists of three rows of pictures depicting seasonal activities of children, and the task imposed upon the examinee involves a knowledge of the two seasons summer and winter and the power to abstract the two of the four scenes that indicate the season.

1. Distinctive Features. There is no time limit; children are not subjected to the emotional strain which a speeding-up process imposes upon the very young. No attempt is made to make sharp differentiation in mentality; its crudeness renders the Test a highly adaptable measure calculated to yield readily local norms. While the norms are but tentative scores,

examination of 895 children has produced a percentile rank table to serve the purpose of comparison. This rank is presented in the form of a table showing " the percentage of 895 pupils whom a child exceeds who makes any given score." Thus a score of 25 was made by 16 of the 895, and this score exceeded the scores made by 95% of the group.

2. Criticism. Criticism centers on choice of cuts, arrangement, and instructions.

Some of the pictures are too vague; others are ambiguous. Pictures at best are poor substitutes for concrete objects, and when line cuts are used the strain on the imagination and memory is severe. A crescent is not a happy presentment of the moon, and an unshaded broken stick may well be mistaken for a hairpin. When children, particularly girls, label this drawing a hairpin, the examiner may well understand the blank look that greets his query " Which two of these three things (broom with broken handle, stick, and shovel handle) belong together? " It is rather unreasonable to expect the examinees to wield a broom with a hairpin. The test based on scenes representing summer and winter would gain in differentiating capacity if the pictures were less fragmentary, less sketchy, and more sharply contrasted. One scene shows a boy enjoying a swim. He is almost entirely submerged under three waves that have the " wave " permanently removed while the one visible hand rests limply on the becalmed surface. Another boy is engaged in the truant pastime of fishing, but the

line he grasps in his lifeless hand shows neither cork, nor fish, nor ripple; there is no action, no dramatic element, no anticipatory attitude.

The arrangement is not of the best. Test 2 should present the top row, the one with which the remaining four rows are compared, in a detachable form or should be printed separately. When the child reaches the lowest row, his eye must traverse the intervening ones, to the distraction of his attention and interest.

The instructions might be altered to meet poverty of vocabulary, characteristic of children born of foreign parents. The kindergartens of many of our cities recruit their children from homes where no English is spoken; and the street has not furnished the associations that supply this deficiency, for parental oversight attaches the child to its home and restricts or circumscribes the environment. Consider these statements: Two of these things belong together; two things in this picture are used in the same way. Show me one thing that is nearest like these two. The word " belong " has a possessive tang, " used," has an indefinite meaning, and " nearest " conveys a special relation.

Nevertheless, this Kindergarten Test possesses real discriminative capacity. Its correlation with teacher judgment is exceptionally high. It may be used with advantage in classifying pupils in three groups, slow, normal, and fast.*

* The "Exercises for Self-Testing and Investigation" which begin on page 132 are planned for use with this chapter also.

Bibliography

Bibliography of Tests, Public School Publishing Co., Bloomington, Illinois

Bibliography of Tests for Use in Schools, World Book

Board of Education, New York City — *Pupils' Progress through the Grades*, Publication No. 6, 1922

Burt, C. — *Mental and Scholastic Tests*, London County Council, London, 1921

Dickson, V. E. — *Mental Tests and the Classroom Teacher*, World Book, 1923

Haggerty, M. E. — "Intelligence and Its Measurement," *Journal of Education*, Vol. 12, pp. 212–216, April, 1921

McCall, W. — *How to Measure in Education*, Macmillan, 1922

Monroe, W. S. — *Measuring the Results of Teaching*, Houghton Mifflin, 1918

Monroe, W. S., De Voss, J. C., and Kelly, F. J. — *Educational Tests and Measurements* (Revised Edition), Houghton Mifflin, 1924

Standardized Tests for Use in High Schools, Public School Publishing Co., Bloomington, Illinois

Thorndike, E. L. — "Intelligence Examinations for College Entrance," *Journal of Educational Research*, Vol. 1, May, 1920

Whipple, G. M. — "The National Intelligence Tests," *Journal of Educational Research*, Vol. 4, pp. 16–31, 1921

Wylie, A. T. — "The Correspondence between the Results of the Illinois Examination and Binet-Simon Scales," *Journal of Educational Research*, Vol. 6, pp. 274–275

Yoakum, C. S., and Yerkes, R. M. — *Army Mental Tests*, Henry Holt, 1920

CHAPTER V

GROUP TESTS OF INTELLIGENCE

Individual tests are superior to group tests. The chief virtue of the former resides in the fact that personal contact enables the examiner to adjust the test to the ever changing moods that often color the examinee's responses; self-consciousness may induce confusional thinking, and overconfidence may betray the judgment into grave mental lapses. The group test places the examiner in the attitude of the physician who administers an anæsthetic without the precaution of keeping a firm hand on the patient's pulse: the examiner is denied the corroborative evidence of imaginative insight. Mind is a highly complicated machine. The slightest emotional disturbance may bring the wheels to a jarring stop, and the immediate presence of the engineer may avert untoward happenings. But time is an inexorable task master. Large school systems cannot afford individual tests; and those charged with testing large school populations have been unable to withstand the steady pressure exerted by school administrators for a rapid survey of new entrants: pupils must be classified as quickly as possible. Therefore the use of the time-saving group test.

Representative Group Tests

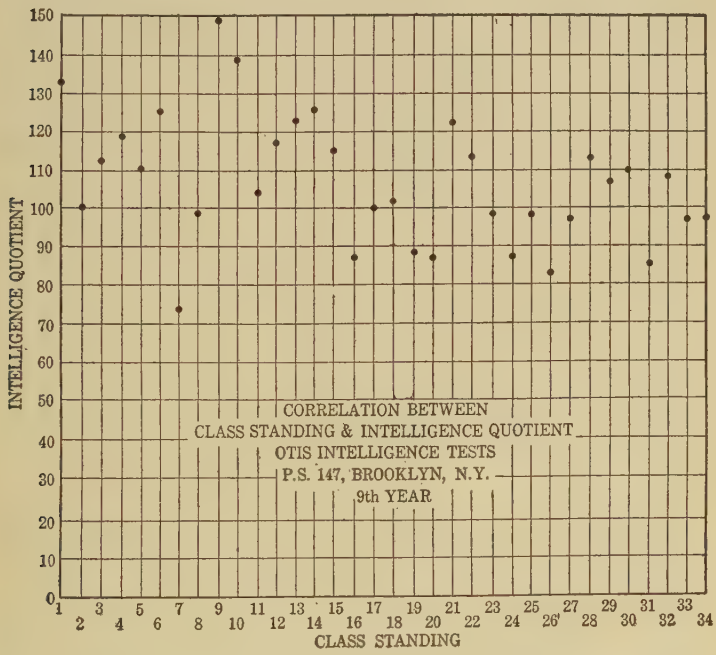
(1) *Otis Group Intelligence Scale* *

1. Purpose. In his introduction to the Manual, Dr. Otis characterizes his Scale as " an instrument for the measurement of native mental ability," which probably comes as near testing raw " brain " power as any system of tests yet devised. He qualifies the statement by remarking that " no one would claim — certainly not the author of this scale — that the method of group testing should take the place of the Binet and other methods of individual examination. The two methods supplement each other."

2. Description. The Scale is designed for both primary and grammar grades. The Primary Examination is for grades I to IV. It consists of eight tests of a nonlinguistic nature. The Advanced Examination is made up of ten tests. It is calculated to test all literates from grade IV to grade XII (elementary school and four years of high school). To prevent coaching, Dr. Otis has devised for each test two alternative forms, known as Form A and Form B. It is permitted to divide each test into two subexaminations, one half at each sitting. In younger pupils this method prevents the ravages of fatigue. For purposes of " rapid survey," abbreviated forms are provided consisting of a nine-, eight-, and seven-test scale. The eliminated tests are

* Copyright, 1922, by World Book Company, publishers, Yonkers-on-Hudson, New York.

weighted, nevertheless, by the use of conversion tables that give final scores in terms of the unabbreviated tests. Thus a ten-test score of 100 is equal to a seven-test score of 66, an eight-test score of 73, and a nine-test score of 84.



3. Distinctive Features. The Scale has been repeatedly revised. The author has spared no pains to include everything that would contribute to its usefulness. In establishing norms, the author has taken careful account of the fact that dull pupils tend to drop out of school. The tables are unusually elabor-

ate. There are grade norms, age norms, and I.Q. equivalents. The Advanced Examination contains many crucial tests, tests that show high correlation with estimated intelligence. Test 4 gives a number of proverbs and twelve statements that explain them in concrete form. The pupil is directed "to find the statement that explains each proverb." The two extra statements tend to defeat chance correspondence and to insure concentration. It seems to favor children who possess (1) a fund of rich social experience and (2) similar vicarious experience gained from wide reading. It is distinctly a test for supernormality. Test 7 is equally important. It is a test of "analogies." The statements are arranged in the form of a proportion, and the examinee is instructed to give the answer to the query "*Clothes* are to *man* as *fur* is to what?" This test also appeals to the very bright child because it makes heavy demands upon one's power of comparison and abstraction.

4. Criticism. While the Primary Examination is fairly free from a linguistic bias, the Advanced Examination presumes a familiarity with the English language. It is therefore not suited to foreign pupils who do not read English.

It is felt that the linguistic element is preponderatingly stressed, and very few investigators have succeeded in clothing the question in a transparent vesture of words. Language plays a leading rôle in tests involving the complexities of a problem. The time con-

sumed in its solution is divided between advancing the stages of reasoning from the recognition of a problem to the verification of the definitive solution, and the penetration of the linguistic shell in which it is inclosed. Of two subjects tested, the one endowed with finer literary discrimination can come more quickly at grips with the problem and can employ the time assigned to greater advantage: their intelligence quotients may show a great disparity and still fail to reflect the actual status of their intelligence. This disability may account for the divergences that exist in the scores obtained in tests of manual dexterity and nonlinguistic tests of intelligence. Poor linguists are afforded the opportunity of exercising their intelligence under conditions in which linguistic difficulty does not retard the reasoning tempo; no verbal obstacles impede the progress of solution by creating a diversion at the moment of its initiation.

(2) *National Intelligence Tests* *

1. Purpose. The purpose of these tests is to adapt the Army Alpha Intelligence Examination to the testing of school children. The aim of Terman and Yerkes, the sponsors, was " to produce in a single pamphlet an arrangement of tests that could be applied to any child in the elementary school who could read well enough to participate in a group examination." The range

* Copyright, 1922, by World Book Company, publishers, Yonkers-on-Hudson, New York.

of difficulty of each test was determined by contriving " that not more than 10 per cent of perfect scores should be made by the average group of eighth-grade pupils and not more than 10 per cent of zero scores should be made by the average group of upper third-grade pupils." This arrangement automatically restricts the scores at the lower and upper limits, and the tests are therefore limited in applicability to grades III to VIII. The tests do not pretend to measure absolute intelligence, but are designed to test the relative order of intelligence obtaining in a group, thus serving as a basis of classification for homogeneous sectioning.

2. Description. The National Intelligence Tests consist of two tests, called Scale A and Scale B. Scale A is more commonly used. It is a battery of five subtests, comprising tests in (1) arithmetical reasoning, (2) sentence completion, (3) logical selection, (4) designating words of same and opposite meanings, and (5) using a symbol digit. " Arithmetical Reasoning " tests the pupil's ability to solve simple problems. " Sentence Completion " is a modified completion test in which the examinee is asked to supply a missing word. " Logical Selection " calls upon the child to decide which things or qualities are essential and which are adventitious. The instructions direct the attention to a sample: Man (body — cane — head — shoes — teeth). " Two of them tell what things a man *always* has. A man does not always have a cane. But he does always have a body and a head; so that

the word *body* has a line under it and the word *head* has a line under it." " Same — Opposite " provides a series of parallel words, with the injunctions to mark with " S " those that are similar and with " D " those that are of opposite signification. " Symbol — Digit " supplies a key with which to mark a series of drawings, all of which are extensively duplicated and scattered promiscuously throughout the page. Test 4 seems to be the crucial test.

3. Distinctive Features. The final selection and arrangement of items for this Test were based on extended research. The inclusion of " power " tests, tests that do not depend on the time element, adds to its value. These tests are so timed that most children finish before the expiration of the allotted time; this neutralizes effectively the disadvantage of combining into one test speed and range of difficulty. The introduction of preliminary exercises tends to allay examination fear. These exercises precede the actual tests, thus familiarizing the children with the nature of the problem; they abate the confusion that the unfamiliar may inject.

4. Criticism. Form A and Form B do not seem to be of equal difficulty. Test 1 of Form A seems to give a greater number of successes than the corresponding test of Form B. This disadvantage is obviated by giving both forms and averaging the result. A further disadvantage of this test lies in the fact that in Test 2, which is a completion test, there are so many possible

correct answers that the scorer must consult a long list in some cases in order to score a single sentence.

(3) *Illinois Examination* *

1. Purpose. The purpose of the Illinois Examination is to test intelligence by gauging the child's ability to "adjust his thinking to new requirements." This power of adjustment is indicated by the sum of the scores on each of the seven tests of which this Test is composed, and in addition by the achievement in silent reading and arithmetic, translated into achievement ages. The two measures are combined into an achievement quotient (A.Q.). The A.Q. formulates the pupil's "capacity" to learn or to do the work of the school.

2. Description. The Illinois Examination consists of three distinct parts, which may be used independently: (1) A Group Intelligence Scale comprising tests of (*a*) analogies, (*b*) arithmetic problems, (*c*) sentence vocabulary, (*d*) substitution, (*e*) verbal ingenuity, (*f*) arithmetical ingenuity, and (*g*) synonyms and antonyms. This is the same for all grades. (2) Monroe's Standardized Reading Tests (Revised). This is composed of Scales I and II. Scale I is for grades III, IV, and V; Scale II, for grades VI, VII, and VIII. (3) A

* Published by the Public School Publishing Company, Bloomington, Illinois.

Note. This Test has been revised by G. M. Whipple and H. D. Whipple to insure "greater accuracy in the directions to the teacher and to the pupils." No revisions have been made in the subject matter of the Scale.

Scale for measuring ability in the operations of arithmetic (Monroe's Standardized General Survey Arithmetic Scales). This is composed of two scales, one for grades III, IV, and V, and the other for grades VI, VII, and VIII. The three scales are bound together in one booklet.

For both reading and arithmetic, the authors have worked out norms in terms of ages. There are arithmetic age norms and reading age norms. These are known as achievement ages. For each age norm there is a corresponding achievement quotient in terms of mental age and achievement age. Thus the median rate score for pupils having a mental age of $11\frac{1}{2}$ years is 145 words per minute; the median score in comprehension for the same group is 11. Therefore a pupil who has a rate score of 145 and a comprehension score of 11 has an achievement age of $11\frac{1}{2}$. If his actual mental age, as shown by the intelligence scales, is also $11\frac{1}{2}$, his A.Q. is found by dividing his achievement age by his mental age, which in this instance is 100. A comprehensive chart furnishes equivalent achievement quotients.

3. Distinctive Features. The use of A.Q.'s dispenses with teacher estimates. It provides the supervisor with norms for classification and the teacher with checks upon the efficiency of her teaching. Another distinctive feature is the ease with which point scores are converted into mental ages. The rule is: Add 43 to the point score and divide by 10.

4. Criticism. The intelligence scale contains no new material. The sentence-vocabulary test is a modified information test, and the substitution test is the symbol-digit test of the National Intelligence Tests. Again, the complementary achievement test is not thoroughgoing enough to indicate scholastic proficiency. Silent reading is a major subject, but it is not an index of all-round linguistic efficiency; composition and, to a lesser degree, spelling and grammar are important requisites; and while the ability in the operations of arithmetic underlies progress in arithmetic, it may exist independently of the ability to apprehend mathematical concepts. In other words, the achievement quotient is but a partial characterization of the child's pedagogical status.

(4) *Myers' Mental Measure* *

1. Purpose. The Myers' Mental Measure purports to test all ages by a brief continuous scale that is independent of school experience. "Each section of the test sets tasks easy and simple enough for the kindergarten child, and at the same time other tasks hard enough for the university student." It aims to aid the administrator (*a*) to classify his children on the basis of native ability and (*b*) to measure the efficiency of his school organization and his teachers by checking the school product with the abilities of the children concerned.

* Published by Newson and Company, New York City.

2. Description. There are four tests. Some of these are given piecemeal, one question at a time, while others are given as a whole. All tests consist of a series of related pictures arranged in rows. Children are directed "to draw some lines and make some marks." Test 1 is a directions test. Test 2 is a picture completion test. The test is fully explained by means of a preliminary demonstration chart, and during the demonstration it is permitted to extend aid to the slow pupils. At its conclusion every participant knows what is missing in the picture of the box, the dish pan, and the shoe. Test 3 requires the recognition of likenesses in a number of pictures arranged in rows. The first row has "two things, just two, alike in some way" and the last row five things. A demonstration introduces this test. Test 3 is a continuation of Test 2; and we may consider the latter as an extensive preparation for the former. Test 4 is also a similarities test, but with alternative selections of pictures. The likenesses are more subtle, and their abstraction calls for a high degree of discrimination. While Test 3 is based on "first common elements," Test 4 is integrated by "second common elements." The common properties are not obvious; they inhere in their use, composition, or structure.

The measure takes the form of raw score with both grade and age norms, and an intelligence ratio. "Intelligence ratio equals raw score divided by chronological age in months." Intelligence ratio, however, "does not mean much above the high school."

3. Distinctive Features. The Myers' Mental Measure is a nonlinguistic test. It is independent of school attainments in the sense that no element of scholastic learning enters directly into the tests. But the enjoyment of educational advantages, with their training of observation and memory, aids materially in obtaining a good score.

The brevity of the test is a virtue that merits commendation. The examiner can test many groups; the examinee is relieved of nervous strain, and his ordeal is over before the onset of fatigue or ennui.

4. Criticism. The claims of the author that a correlation of 80 obtains with the Stanford-Binet is borne out by an examination of " about 300 school children " with 81 for the first grade and 68 for the eighth grade — a rather modest corroboration. The four tests are really two tests — directions and similarities; and these can test only a phase of intelligence. These reservations do not, however, detract from its usefulness as a supplementary diagnosis of intelligence provided the raw score only is used. The intelligence ratio is but a species of statistical manipulation that is more ingenious than valid.

(5) *Dearborn Group Tests of Intelligence (Revised)* *

1. Purpose. The purpose of the Dearborn Group Tests of Intelligence is to determine the degree of intelligence in terms of the mental ages, and the intelligence

* Published by J. B. Lippincott Company, Philadelphia.

quotient. According to Dearborn the limit of mental growth is reached at 14, though in computing I.Q. the author recommends the use of $14\frac{1}{2}$ as the denominator — a concession to the psychologists who consider 16 as the upper limit of mental development.

2. Description. The tests are arranged in two series. Series 1 is composed of General Examinations A and B, and is suitable for grades I to III, inclusive. Series 2 is designed for grades IV to XII, inclusive, and contains two General Examinations, B and C. The two parts are to be given on separate days. Either form of the General Examinations may be used as an abbreviated test, and separate norms are provided. Series 2 has been latterly revised, and needed norms have been worked out. General Examination C is a four-test folder. Test 1 consists of picture sequences arranged in rows, each row telling, in deranged sentence sequence, a complete story. The examinee is directed to number in proper order four pictures showing, for instance, a postman delivering a letter, a boy reading a letter, and a girl mailing a letter, followed by one portraying a girl writing a letter. Test 2 is one of "word sequences." This is a verbal counterpart of Test 1. The words "dinner, supper, breakfast," for example, must be numbered to correspond to the chronological order of the occurrence of action to which they refer. "Breakfast," of course, should be numbered 1. Test 3 is known as "Form Completion." This test contains four shaded triangles numbered 1, 2, 3, and 4. Below

these are blocks with blank and shaded spaces. The examinee is invited to fit the numbered triangles into the blank spaces by blocking in the figures with dotted lines. This is a judgment test. Test 4 calls for " opposite completion." " One or more words have been left out of the following sentences. Complete each sentence by writing in the blank space a word which means just the opposite of the word which is underlined." The word underlined is the logical opposite of the one to be supplied, thus producing an intelligible sentence.

3. Distinctive Features. The Dearborn Tests strive to preserve the balance between linguistic and non-linguistic tests. Grade norms are not stressed, an omission that deserves commendation as children of widely varying ages are usually found in the same grades. Of especial administrative significance are the two interpretations of the I.Q's. The author states that: " In doubtful and borderline cases and especially when the intelligence quotients are below 80 or over 120, these tests should, when possible, be supplemented by individual mental tests and standard tests of school accomplishment."

4. Criticism. If the discriminative capacity of these tests is of such dependability as to warrant their use for administrative purposes, Dr. Dearborn has given us a very valuable measure of intelligence. Coming from a psychologist of his standing, his conclusions merit serious consideration.

(6) *Haggerty Intelligence Examination* *

1. Purpose. The aim of the Haggerty Intelligence Examination is to discover " that capacity which enables a pupil to profit by the work of a good school, to recognize the meaning of words and sentences, to remember important facts observed and learned, to discriminate essential from unessential materials, and to infer correct conclusions from presented data."

2. Description. The Intelligence Examination is composed of two scales, Delta 1 for primary children and Delta 2 for the upper grades.

Delta 1 consists of six tests, each preceded by a preliminary " fore-exercise." " The fore-exercises are printed as Exercises 1, 3, 5, 7, 9, and 11; the tests proper, as Exercises 2, 4, 6, 8, 10, and 12." Exercise 2 is called " Oral Directions." It directs the child to draw certain marks and lines in connection with pictures arranged in rows. Exercise 4 involves " copying designs." Exercise 6 is a picture completion test. Exercise 8 is devoted to " picture comparison." Pictures are arranged in two's, and the examinee must indicate those that are similar with " S " and those that are dissimilar with " D." Exercise 10 is known as " Symbol-Digit." Here, instead of marking similarities, likenesses are designated by numbers. Both the Dearborn and the National Tests include such tests. Exercise 12

* Copyright, 1922, by World Book Company, publishers, Yonkers-on-Hudson, New York.

seeks out similarities and differences in words. It is called " Word Comparison."

Delta 2 is designed for grades III to IX and nowise differs from other tests of intelligence. It is composed of seven exercises, sentence reading, a "yes or no" test, arithmetical problems, picture completion, synonym-antonym tests, and the use of practical judgment (as " Why do we use stones? " with three possible reasons), and a test of common sense. Instead of " fore-exercises " elucidating directions and samples are at the head of each exercise.

3. Distinctive Features. By administering the Intelligence Examination in combination with the Achievement Examination in Reading, the author endeavors to obtain a cross section of the pupil's potentialities. It is not essential to give the reading test owing to the fact that Exercises 1 and 2 of Delta 2, tests in reading and arithmetic respectively, have been standardized individually for all grades. Delta 2 may be considered tests of both intelligence and achievement.

The author is to be commended for extreme moderation of language. He is fully aware of the inconclusiveness of any one test or the use of an individual test. Especially illuminating is his admonition to employ supplementary aids under the guidance of "common sense." The administrator " must use the results of his test to supplement, correct, or to reënforce his ordinary powers of observation. Only so will the tests justify themselves."

4. Criticism. The scoring of the answers in the nonverbal tests in both Delta 1 and Delta 2 is not sufficiently objective. The younger children do not possess enough precision in motor coördination to draw the marks and lines exactly as directed. Many lines have a tendency to pursue a tortuous path, touching other objects in the way. A cross, for instance, may be prolonged to extend to other pictures.

(7) *Detroit First-Grade Intelligence Test* *

1. Purpose. The author, Anna M. Engle, has devised the Detroit First-Grade Intelligence Test with a view to aiding administrators to classify children entering the first grade. It does not attempt to establish mental age or I.Q.; it merely groups children according to the magnitude of the scores attained.

2. Description. This intelligence test is presented as a game in which no more than twelve examinees participate. It is made up of ten tests composed of pictures. The participants mark with a cross those designated by the examiner. Each test contains five parts or rows arranged in order of difficulty. Test 1 is an information test. One of the rows displays a picture of a spoon, scissors, hammer, and knife. The examiner's instructions are " to mark two things we cut with." Test 2 is one of similarities. A series of related drawings (five and one sample in each series) is pre-

* Copyright, 1922, by World Book Company, publishers, Yonkers-on-Hudson, New York.

sented. The examinee is directed to mark the first one and then mark another one just like it. The points of likeness grow increasingly subtle and make great demands on the power of observation. Test 3 is a direction test. In the first row one object must be marked; in the fifth row five things must be crossed out. This test measures the memory span, as the order must be executed after listening to the complete instructions. Test 4 is a test of discrimination functioning in the detection of absurdities. "There is something wrong with every one. What is the matter with the dog? Yes, he has two tails. Mark the wrong one." Test 5 is another test of discrimination; magnitude is the medium for its manifestation. The examinee cannot comply with the examiner's request to "mark the biggest block" unless he can discern small differences in area. Test 6 is one of logical connection. One of the directions reads, "Mark the thing you use with a knife." This "thing" is companioned by a cup, spoon, pot, and fork. Test 7 is a numerical completion test. A series of balls is shown in various alignments from which one ball is missing. The alignment must be completed by the examinee. A sense of symmetry should prove of great aid. Test 8 tests the child's motor coördination and observation. The examinee is called upon to complete a picture so as to make it look like the accompanying sample. Test 9 appraises the child's ability to count. "See the stars? Mark 13 of the stars." Test 10 follows the Myers

Mental Measure. It is a directions test in which the examinee is asked to connect pictures by a line that must pass over or under adjacent ones.

3. Distinctive Features. It is nonlinguistic and does not require the child to do much. The *Examiner's Guide* is succinct and brief and contains both directions and test. The norms lend themselves to ready adaptation to meet local conditions. The author gives the true median and the 25th, 50th, and 75th percentiles. These are cast into a system of ratings.

E	0 to 12 points
D	13 to 17 points
C	18 to 23 points
C⁻	24 to 30 points
C⁺	31 to 35 points
B	36 to 39 points
A	40 to 50 points

E and D are the " slow-moving groups, to be instructed in minimum essentials only." C, C⁻, and C⁺ are " average rate groups, a group where course of study is designed for pupils of average ability." A and B are " fast-moving groups to be given enriched course of study."

4. Criticism. The children for whom the test is designed are too immature to work in groups and are lacking in the concentration that mass teaching demands. Nor do they possess sufficient self-confidence to follow directions without personal guidance.

They are rather heartened by expressions of intimacy, the pointing finger, the cheering compliment, the steadying voice — attentions that only the closest proximity vitalizes. The author's monition to help, if necessary, with the first line of each test converts this group test into an individual test; and as an individual test it has real merit. But it is, as yet, too crude to justify such fine grouping. E and D might better be combined into one, C^-, C^+, and C into another, and A and B into a third; C, B, and A are sufficiently classificatory for administrative purposes.

(8) *Otis Self-Administering Tests of Mental Ability — Intermediate and Higher Examinations* *

1. Purpose. The tests are designed to measure the mental ability of all persons able to read. The Intermediate Examination is designed for grades IV to IX and the Higher Examination for grades VII to XII and college students. Each examination is furnished in two forms — Form A and Form B. The tests are in the nature of a revision of the Otis Group Intelligence Scale — Advanced Examination.

2. Description. Each of the two tests consists of 75 items arranged in a single list. The author seems to have used the crucial tests of the Advanced Examination as a standard. Many questions are borrowed;

* Copyright, 1922, by World Book Company, publishers, Yonkers-on-Hudson, New York.

others are frank adaptations. The questions may be classified into types:

(*a*) *Discrimination:*

A contest always has (?).

(1) an umpire, (2) opponents, (3) spectators, (4) applause, (5) victory . . . . . ( )

(The number and not the correct word is to be inserted in the parenthesis.)

(*b*) *Mathematical Reasoning:*

If a wire 20 inches long is to be cut so that one piece is $\frac{2}{3}$ as long as the other piece, how long must the shorter piece be? . . . . . ( )

(*c*) *Analogies:*

What is related to a cube in the same way in which a circle is related to a square? . . . . ( )

(*d*) *Ingenuity:*

Suppose the first and second letters in the word *constitutional* were interchanged, also the third and fourth letters, the fifth and sixth, etc. Print the letter that would then be the twelfth letter counting to the right. (.)

(*e*) *Judgment:*

A man who is averse to change and progress is said to be ( ).

(1) democratic, (2) radical, (3) conservative, (4) anarchistic, (5) liberal

(*f*) *Opposites:*

The opposite of *treacherous* is ( ).

(1) friendly, (2) brave, (3) wise, (4) cowardly, (5) loyal

3. Distinctive Features. The Test is very easy to administer. The examiner merely hands out papers, directs the examinees to "do what it tells you to do," accords them an interval in which to familiarize themselves with the instructions, announces the time — 30 minutes for a thorough examination and 20 minutes for a rapid survey — and collects the papers at the expiration of the time limit.

It correlates highly with teachers' estimates and may be used with gratifying results in junior high schools. It singles out, with fair success, the student who is capable of doing abstract thinking, involving the use of language and mathematics.

4. Criticism. The twenty-minute period is entirely inadequate for individual diagnosis, and is not recommended by the author of the Test except for rapid survey purposes. Many of the questions make severe demands upon the examinee's concentration. Other questions entail preliminary rearrangement of the data before the solution may be attempted. The last twenty questions tax the most ingenious. The predictive capacity of this Test might be enhanced if no time limit were set. The Higher Examination is not a "speed" and "time" test; it is rather a test of "power."

(9) *Trabue Mentimeters (School Group 2 A)* *

1. Purpose. The name is self-explanatory; they are designed to test and classify "intellectual capacity" of pupils. Of the thirty mentimeters published in *Measure Your Mind*, School Group 2 A contains the eight which in the opinion of Trabue are suited to be used "as a group test of the intellectual abilities of children and adults of all ages, being especially selected for the classification of students according to their relative abilities to succeed in general academic work."

2. Description. Thirty subtests, each independent of the other *administratively*, constitute the group of tests designed to render mind testing intelligible to the layman. The composition of School Group 2 A is as follows:

Number of Original Mentimeters	Title
2	Pictorial Absurdities
3	Maze Threading
5	Dividing Geometrical Figures
16	Naming Opposites
18	General Information
20	Reading: Directions
23	Completion of Sentences
28	Arithmetic Reasoning

Numbers 2, 16, 18, 20, 23, and 28 afford little that is new to the reader. They differ from similar standard tests in point of time allotment and in number of items in each mentimeter. Tests consume from 180 to 600

* Copyright by Doubleday, Page and Company.

seconds; the number of items varies from 12 to 40, the mean being 23.

Mentimeter 3 recalls the Porteus Maze. It is called "Maze Threading." There are twenty mazes of progressive difficulty. The examinee is called upon to thread a maze in direct line from the point of ingress to the point of egress, arrows indicating entrance and exit. The directions follow those standardized by Porteus.

Mentimeter 5 is known as "Dividing Geometrical Figures." The pupil's attention is called to the explanatory exercise, which consists of "a large figure at the right (a circle) and two or three small ones at the left (arcs of the same circle). When properly put together, the small figures would make the large figure." A dotted line is to be used to indicate the superposition.

3. Distinctive Features. An elaborate manual supplies both directions and psychological justification. It is known as *Measure Your Mind.* The authors are Stockbridge and Trabue. The *Mentimeter Record Booklet* provides space for the recording of the results of fifty examinations, under the rubrics of score, per cent lower rank, and frequency of occurrence. A child — one of a group of 40 — makes a score of 22. Suppose there are twenty other individuals in this group or 50% who make lower scores, and that three others make identical scores when tested by the same mentimeter. This examinee's status would be summarized thus: J. D. — Age 15; score in Mentimeter 22; lower 50; rank 18.

In other words, J. D. with a score of 22 surpasses 50% of his group and ranks as eighteenth in a class of 40.

Both grade norms and age norms are provided. The grades range from I to XII; no age norms are supplied for the first six years, and the highest age is sixteen. Mentimeter (age) index gives the mental age; Mentimeter quotients take the place of the Binet I.Q.

4. Criticism. Trabue makes a clear distinction between the Binet I.Q. and the Mentimeter quotient. He is not led astray by surface indications. He maintains a judicial attitude of cold detachment and resists the temptation to substitute statistical plausibility for scientific precision. His judgment of the value of group tests is straightforward and unambiguous; he wants it clearly understood that "the Mentimeters and other group tests are not as reliable as is the Standard Revision of the Binet Tests, and that a deviation from the standard in any group test is therefore not as significant as a corresponding deviation in the more consistent individual psychological examinations." No one can reasonably object to window dressing an article.

Exercises for Self-Testing and Investigation

The items under A, B, and C should be matched so as to construct a series of ten statements embodying the author or authors, the name of the test credited to him or them, and a sentence or phrase characterizing the general tenor of the test.

A

1. Kuhlmann
2. Pintner-Paterson

3. Yerkes, Bingham, Goddard, Haines, Whipple, Terman, and Wells
4. Yerkes, Bridges, and Hardwick
5. Otis
6. National Research Council
7. Monroe and Buckingham
8. Porteus
9. Thorndike
10. Haggerty

B

1. Group Intelligence Scale, Advanced Examination
2. National Intelligence Tests
3. Revision of Binet
4. Performance Scale
5. Delta 2
6. Intelligence Examination
7. The Point Scale of Intelligence
8. The Illinois Examination I and II
9. Army Alpha Group Intelligence Test
10. Maze Scale

C

1. The addition of tests for ages below the third year
2. Contain a picture-completion subtest
3. Used successfully with foreign children
4. Devised for use in high school
5. Is a combination test (intelligence and achievement)
6. Gives partial credits instead of using "all-or-none" method of the Binet tests
7. Contains a subtest built on proverbs
8. Subtests preceded by an "exercise"
9. All tests carried out by tracing with a pencil on a printed maze
10. Prototype of many group intelligence tests

Bibliography

Binet, A., and Simon, T. — *The Development of Intelligence in Children*, Williams and Wilkins, Baltimore, 1916

—— *A Method of Measuring the Development of Intelligence in Young Children*, Chicago Medical Book Co., 1915

Brigham, C. C. — "An Experimental Critique of the Binet-Simon Scale," *Journal of Educational Psychology*, pp. 439–448, 1914

Burt, C. — *Mental and Scholastic Tests*, London County Council, London, 1921

Dickson, V. E. — *Mental Tests and the Classroom Teacher*, World Book, 1923

Goddard, H. H. — *The Binet-Simon Measuring Scale for Intelligence* (Revised), Training School Bulletin, Vol. 8, pp. 56–62, 1911

Gordon, K. — "Some Retests with the Stanford-Binet Scale," *Journal of Educational Psychology*, Vol. 13, pp. 363–365, September, 1922

Haggerty, M. E. — "Intelligence and Its Measurement," *Journal of Educational Psychology*, Vol. 12, pp. 212–216, April, 1921

Healy, W. — *The Individual Delinquent*, Little, Brown, 1915

Kuhlmann, F. — "Binet-Simon's System for Measuring the Intelligence of Children," *Journal of Psycho-Asthenics*, Vol. 15, pp. 78–97, 1911

—— "A Reply to Dr. L. P. Ayres's Criticism of the Binet and Simon System for Measuring Intelligence of Children," *Journal of Psycho-Asthenics*, Vol. 16, pp. 58–67, 1911

—— "A Revision of the Binet-Simon System for Measuring," *Journal of Psycho-Asthenics*, September, 1912

McCall, W. A. — *How to Measure in Education*, Macmillan, 1922

Monroe, W. S. — *Measuring the Results of Teaching*, Houghton Mifflin, 1918

Monroe, W. S., DeVoss, J. C., and Kelly, F. J. — *Educational Tests and Measurements* (Revised), Houghton Mifflin, 1924

Thorndike, E. L. — "Intelligence Tests," *Seven Famous Reprints from New Britannica, Thirteenth Edition*, Encyclopædia Britannica Co., 1926

CHAPTER VI

ACHIEVEMENT TESTS AND SCALES

A. As Objective Measurements

Scholastic proficiency is one of the facts of school success. Achievement tests and scales attempt to evaluate it through the instrumentality of objective measurements of scientific precision. Tests and scales have been constructed along lines analogous to those underlying physical measurements. The results have been approximative and more or less arbitrary because educational products are rarely one-dimensional and because tests are not exact measuring rods. A board possesses many properties; the builder must measure them somehow. He wants to know how long it is, how much weight it can support, how durable it is, and how much it costs. For each of these qualities there exists a measuring instrument: the yardstick, the machine for gauging tensile strength, empirical knowledge gained through continuous association with this kind of wood, and the dollar. Excepting the measure of durability, all these are scientific; they are practically constant; the units are of equal value; there is a zero point; they measure one quality; the measure is the highest point reached; and they are

thoroughly objective. To ascertain the lasting qualities of the board, one must weigh subjectively certain factors that make for disintegration, such as (1) degree of exposure to elements, (2) use to which it may be put, (3) use or nonuse of preservatives, and (4) attacks of insects. The builder might attain scientific exactness were he to devise a scale which assigns to each cause a maximum, say, of five marks, giving a total of twenty points under the most favorable conditions for preservation. Such a scale would be more exact than a composite judgment but would fall short of exactness. It lacks the precision and infallibility of the yardstick. Observe that the builder has at his disposal three instruments of precision, and that the information he seeks may be gained through four distinct sources and not through one composite measure.

Now consider the evaluation of an educational product, oral reading. To teach efficiently, the teacher must know (1) how fast the pupil reads, (2) how excellently he reads, (3) how much he reads in a given time, and (4) how difficult a passage he can read. Four factors are inextricably bound up in this product: time, quality, quantity, difficulty. Unlike the builder, the teacher cannot separate them for individual examination. He is also forced to deal with four variables. Nor can he altogether eliminate any three so as to segregate them one at a time. He must, in his testing, employ some reading material. Suppose he wishes to test the time element. Any reading material is bound to be on some

level of difficulty, of some quantity. It must be read with some degree of excellence. The nearest approach to the condition of a single variable is that known as testing for the " dominant variable." It has been proposed that reading be tested in a two-dimensional plane, quantity and quality; that " quality " be reduced to a level " just well enough to recognize that it says something or just well enough to be able to do what it says to do," that " quantity " might be eliminated by directing pupils to check each word they recognize. The Thorndike Visual Vocabulary Scales test quality chiefly by requiring the pupil to " write the letter *F* under every word that means a *flower*, the letter *A* under every word that means an *animal*, " and so on.

The complications that confront testing for quantity and quality are multiplied in assessing the degree of difficulty. Not only is there a compromise made with the law of the single variable, but the method of ascertaining the degree of difficulty of a paragraph or selection is characterized by arbitrariness. Such tests are constructed on purely statistical foundations. This is not the only departure from scientific accuracy. We shall see that educational tests violate, to a greater or lesser degree, the principles of scientific measurement as embodied in the yardstick; that they ignore the principle of a single variable; that they contain artificial zero points; that the steps are not equally distant or of equal value; but that, in spite of their serious shortcomings, they are superior to the traditional test,

vitiated as it often is by the vagaries of personal judgment.

What School Products Are Tested

Analysis of the school subjects discloses elements or units of learning as follows: (1) Facts of arithmetic, history, and geography. (2) "Bonds." These are partly synonymous with conventional and tool facts, but refer specifically to discrete associations, such as multiplication tables and mathematical combinations, which, though similar in composition, must be learned separately, owing to their nontransference quality. (3) Processes in arithmetic, algebra, and science. (4) Skills such as writing, drawing, and oral reading. (5) Simple products such as spelling. (6) Complex products such as composition and silent reading. (7) Application of facts to new situations involving the use of judgment and reasoning, comprehended under the word "power." The latter designation places these tests in direct antithesis to the preceding ones, which are known comprehensively as "knowledge" tests. Recent tests in civics aim to discover the pupil's ability to apply knowledge to new situations. Examples of power tests are problems in arithmetic, judgment tests in history and geography, and silent reading. The classification includes some duplication, but they are in some of their aspects sufficiently distinct to warrant their being placed in separate categories.

B. Tests and Scales

The tests devised to evaluate these school products are known variously as tests or scales. Both are measuring instruments. McCall differentiates them on the basis of use. When the instrument is placed in the hands of the pupils and his reactions noted, we have a testing instrument. Thus Thorndike's Visual Vocabulary Scale A–2 is, according to him, not a scale but a test. So is Woody's Addition Scale. When the pupil's composition is compared with standardized samples and a valuation assigned to it corresponding to the sample it resembles the most, we have a scale, a scoring instrument. He distinguishes them as product scales and performance scales. "The performance scales are placed in the hands of the pupils in order to test them. They are real testing instruments and not scoring instruments. . . . Collection of the pupils' composition specimens is the composition test. The composition scale is only the scoring instrument." In actual practice, however, the words "scales" and "tests" are used interchangeably. If any distinction obtains, it is, for the class teacher at least, purely academic.

(1) *The Construction of Tests and Scales*

In the construction of these instruments, the method pursued has varied from chance trial and error to meticulous statistical validation. Indeed some of the tests are merely standardized traditional tests. Those de-

visers of tests who aspire to scientific accuracy have followed these steps:

1. The school product to be measured is thoroughly analyzed. Its component elements are scrutinized painstakingly in order to ascertain which are salient. Those elements are rejected which do not contribute to a progression from simple to more complex. The adventitious and irrelevant are ruled out. Dr. Ayres was faced with the problem of selecting those spelling words that would prove acceptable to educators. He decided that only commonly used words were to be included in the Ayres's Measuring Scale for Spelling. He therefore addressed himself to the stupendous task of examining letters, newspapers, and standard literature. He amassed a total of 368,000 words, written by 2500 persons. He chose the commonest 1000 words, which he divided into fifty lists of twenty words each.

2. The material so obtained is tried out on selected groups and unselected groups. These are divided equally between individuals who, in this subject, are decidedly bright or hopelessly incompetent. The tentative test is summarily condemned if it does not decisively discriminate between the two types. In the event of success, the test is given to unselected groups composed of pupils from two successive grades, in order to find out whether there is progress from grade to grade and whether the scores are distributed according to the normal curve of probability. The last precaution is dictated by the need for an instrument

that will segregate not only the dull and the bright but the great array of talent lying between these extremes.

3. Along with the preliminary try-out of the test, provision is made for a tentative formulation of procedure. This includes manner of administering the test and instructions to pupils. Although instructions may appear foolproof, yet there may lurk within an innocuous word or phrase a wealth of misunderstanding, the removal of which may involve the examiner in endless verbiage. Not infrequently directions may err on the side of comprehensiveness. The laudable purpose of providing for every reasonable contingency may tend to distract the attention of the examinee who, intent on following instructions implicitly, may revert too often to them to the detriment of the test. The procedure, then, must be studied with the same care as that bestowed on the test.

4. The pupil's answers are marked. At this point questions that show a wide variety of answers are studied for (1) signs of ambiguity, (2) indefiniteness, or (3) causes provocative of divergent responses. Questions the answers to which are difficult to score are eliminated.

5. This step inaugurates the statistical manipulation of the answers. The results are tabulated by each question for each pupil.

6. This table furnishes the information as to what per cent of the pupils tested answered each question.

7. The table is arranged in a series from low to high.

This is the first intimation the examiner has of the range of difficulties inherent in the questions.

Tests of quantity and tests of quality diverge markedly at this point. The deviser of the former type of test proceeds (1) to rearrange the questions in order of indicated difficulty, (2) to formulate norms, (3) to establish norms, (4) to formulate scoring instructions to try them on another group, and (5) to establish the validity of the test (*a*) by means of correlation with a test of known validity, (*b*) by comparison with the pooled judgment of many teachers competent to pass upon the group examined, and (*c*) by its correspondence with the normal curve of distribution.

On the other hand, a test of quality or difficulty must obtain a decision upon the degree of difficulty obtaining throughout the test. The constructor of such a test must provide such tangible proof of difference that a teacher using this test fifty years hence will be enabled " to verify the accuracy of the amounts of difference."

1. Quality Scales. Experimenters have attempted to solve this problem in different ways. Hillegas, the author of the Hillegas Composition Scale, decided that the unit of measuring the quality of the ten samples constituting the Scale was " such a difference as could be noticed by exactly three-fourths of the 202 judges " to whom the samples were submitted for placement. The validity of such judgment is based on the assumption that equally often noticed differences are equal

unless always noticed or never noticed; and 75% was chosen as the maximum on another assumption, that, given a fairly large number of trained judges, say, several hundred, it may be reasonably assumed that, when three-fourths agree unconditionally upon the merit of a sample, the agreement would hold for an indefinite number of equally competent judges. Since the personal equation is never free from bias or predilection, 75% is near enough to unanimity to satisfy the most exacting. Besides unanimity in terms of being " always noticed or never noticed " possesses no discriminative value, because several causes may contribute to create a situation of extremes in quality: it does not indicate the *amount* of equality present or absent. Briefly, the 75% point is selected because it is midway between absolute certainty (100%) and pure chance. With only two possible judgments, 50% of the judgments would be correct by pure guesses. The tendency is then " to regard a difference in amount of the thing in question which can be distinguished correctly by 75% of the judges or judgments concerning it, as the smallest psychological unit that can be used with reasonable certainty."

Ayres's Handwriting Scale evaluates one dominant quality, legibility. In the construction of this Scale, Ayres proceeded as follows: He collected 1578 samples of handwriting from children of the upper elementary grades of 40 school systems in 38 different cities. Since legibility was the quality to be measured and since

time consumed in deciphering the writing was to serve as a measure of legibility, the samples were composed of words " thrown out of context." Such content would preclude the possibility of increasing the judges' speed of reading through memorization or through the aid lent by suggestiveness of consecutive sentences. Every sample was read by ten assistants who, stop watch in hand, computed the average number of words per minute. The samples were also classified arbitrarily into five classes — vertical, medium slant, extreme slant, back hand, and mixed. Owing to the small number of samples in the " back hand " and " mixed " classes, they were left out of the final scale.

When these dual classifications were made — seriatim according to time consumed and placement in the three " slant " classes — the sample which stood midmost in the distribution for speed was chosen as the central point. This sample was marked 175.7 because it was read at the rate of 175.7 words per minute. Its central position was owing to the fact that one-half of the samples were read at a slower rate and the other half at a more rapid rate. The entire series was then considered as 100, and this sample was therefore given a rating of 50. Other samples were picked out in an ascending series of 10, 20, 30, 40, 50, 60, 70, and 80. They were rated as to speed as 130.2, 149.4, 163.5, 175.7, 186.1, 195.8, 202.9, and 209.6. The final scale contained these eight groups, each group containing three types of writing : vertical, slant, extreme slant.

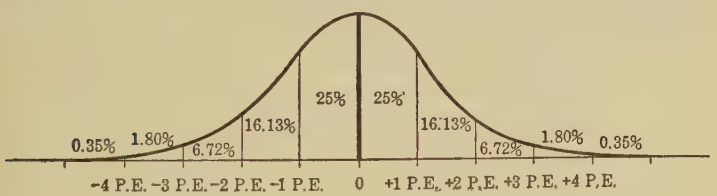
It will be noticed that this scale has no zero point. Neither are the steps of equal value. The inequality of steps is due to the nature of the material. The time consumed in reading samples decreased disproportionately as one approached the upper limits of the scale; the interval between the samples designated 149.4 and 163.5 is longer than that which separates samples 202.9 and 209.6. The equality that does obtain between the steps of this scale is not absolute, but relative; the qualities 60 and 40 (to use the scale designation) are equally distant from quality 50, while qualities 10 and 30 occupy the same relative position with regard to quality 20. It must not be understood, however, that the interval between qualities 20 and 30 is the same numerically as that between qualities 70 and 80.

2. Difficulty Scales. Educational statisticians have sought to attain scientific exactness by eliminating the objectionable features of the quality tests. They have, therefore, devised a method of scaling that would produce a measure having greater constancy, steps of equal intervals, and a zero point of lesser arbitrariness, for they have felt that judges are the product of their environment and judgment is a progressive quality. What might be considered good quality in one generation or even decade might be condemned in the next; and the same examiner may alter his judgment within a short space of time. Besides, judgment is a precarious base on which to build a scientific measuring rod.

The measure whose use they advocate is the probable

error (P.E.) and the standard deviation (S.D.). They are applicable to both quality tests (like a handwriting scale) and difficulty tests (like a problem test). The P.E. measures a smaller interval than the S.D., but their use has been confined to difficulty scales, using the word " scale " generically. This method involves certain assumptions:

The distribution of the trait or the product to be measured follows the bell-shaped curve of normal distribution. True, improved method of teaching



may raise the level of performance, but the pupils' relative standings in a large sampling of the school population will remain the same; there will be just so many very low scores, so many average scores, and so many superior scores.

The dispersion of the group is a measure of difference. In a normal curve of distribution, the base line may be cut up into equal units. At the points of dissection, perpendiculars may be erected. The surface bounded by them will show the number of cases included. Since this is a curve and not a rectangle, the base line will have to be unequal to include proportional areas. The figure on this page illustrates the method.

Along this line, the samples are located according to their difficulty. They are distributed about the median. Thus a problem solved by exactly half the group will be located at the median, or mid-point; a problem solved by more than half would be placed below (at the left of the median); a problem solved by less than half, above (at the right of the median).

Distances along the base line are commonly expressed in terms of the P.E. The P.E. (in the probability curve) "is equal to such a distance on the scale that, when laid off on each side of the average, it will include one-half of the measures," that is, the middle 50% of the group. The probable error expresses the number of units by which the middle 50% of the group deviates from the median. Suppose that the median value of speed in writing is 60 words per minute. A P.E. of 1.5 would signify that 50 out of 100 pupils selected at random from those attending the same grade (if the test is a grade scale) will in all probability write 65 + 1.5 (66.5) words per minute or 65 − 1.5 (63.5) words per minute. The amount of dispersion here used, 1.5, is known as 1 P.E.; *i.e.*, 66.5 words per minute is + 1 P.E. from the median.

Theoretically the base line of the normal curve never touches the curve, but it is "practically" justifiable to break off the base line of the curve at ± 4 P.E. By the breaking of the base line of the curve at ± 4 P.E., 0.35 of 1% of the cases on each side of the average are neglected, that is, in all 0.7 of 1% of the cases.

In any chance distribution of marks 50% of the cases would fall within ±1 P.E. of the median; −1 P.E. or + 1 P.E. distance from the median would embrace 25% of the cases, tapering off in a fixed percentage from the median.

An inferior pupil who is surpassed by 99.5% of a large unselected group of children would be located on the base line of the normal curve at point − 3.82 P.E. below the average. Frequently such a pupil is spoken of as having a percentile of 0.5. A pupil with a percentile of 10 (that is, one who is surpassed by 90% of an unselected group) would be located on the base line of the curve at a point − 1.90 P.E. below the average. Likewise a pupil with a percentile of 25 would be located at a point − 1 P.E. below the average; a pupil with a percentile of 50 would be located exactly at the average; a pupil with a percentile of 60 would be located at a point + 0.38 P.E. above the average; a pupil with a percentile of 75 would be located at point + 1 P.E. above the average; a pupil with a percentile of 90, at + 1.90 P.E.; and a pupil with a percentile of 99.5 at + 3.82.

Having postulated the constancy of the P.E.'s in any chance distribution, the test maker reverts to percentile rank of each question or problem. By referring to table of P.E. values, he assigns to each percentile rank its P.E. location. Thus problem A, which has been solved correctly by 6% of the group, is given a P.E. value of + 2.30; problem B in like manner is given

a P.E. value of 2.13; problem C a P.E. value of 1.90; and so on to the end of the list. All these values are listed in order of P.E. valuation. However, only those are included in the test which are equally spaced along the base line, discarding the plus and minus signs (see figure on this page) of the P.E.'s and expressing the location in a progressive line beginning with the zero point, this point being the example which very few

pupils could answer correctly. This base line has a length of about 8 P.E. (4 minus and 4 plus). The lowest difference deemed significant is .60 P.E.'s. Therefore, the problems might be located as follows:

Problem	A	C	G	H	. . . .	Z
P. E.	− 3.45	− 2.80	− 2.24	− 1.64	. . . .	+ 3.04
Per cent solving it correctly	1.0	2.5	6.5	13.5	. . . .	98

The final step is taken when the problems or questions are numbered consecutively in order evolved through P.E. locations.

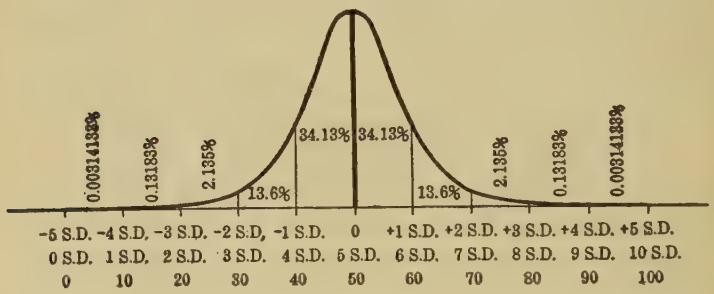
McCall makes use of another measure of dispersion, the S.D. 1 S.D. above and below the mean (average) includes about 66% of the distribution. The technique employed by McCall is essentially the same used in converting a percentile rank into P.E. locations. He does use, however, a sampling of twelve-year-old children, 500 in number, claiming that greater constancy accrues therefrom. His tabulation of data follows:

1. Total number of questions correct
2. Number of twelve-year-old pupils having such numbers correct
3. Number of pupils exceeding *plus* half those reaching each total number of questions correct

Thus one pupil out of a total of 500 answered 33 questions correctly. The number of pupils exceeding 33 was 0; half those reaching 33 was .5. Their sum is .5. This constitutes 1% of the sampling. By means of a conversion table the per cent values are changed into S.D. locations. The plus and minus signs are discarded, thus giving a range of 10 S.D.'s. The zero point is 5 S.D. below the mean. A per cent of 1 yields a location value of 7.3 S.D.'s. This is multiplied by 10 to eliminate decimals and is expressed as 73 T. The median ability of twelve-year-olds is 50 T. Expressed in terms of minus and plus S.D.'s, 7.3 S.D. may be written as + 2.3 S.D. above the mean (see figure on the next page).

The scale for twelve-year-olds in terms of their S.D.

variability is called a T score, in honor of Terman and Thorndike. McCall supplements this T score by a B score (in honor of Binet) by establishing corrections for ages above and below twelve years. The method is rather complicated and need not detain us. According to the B score, a seven-year-old pupil who answers correctly 7 of the questions standardized on twelve-year-old pupils would receive a score of 32, corresponding to the value of the T score, and a correction



of 34 to correspond to his age. His B score would thus read 66.

Both P.E. and S.D. evaluations of difficulty involve a great deal of work. McCall considers the labor well spent because T scores possess:

1. Definiteness and convenience of reference point
2. Equality of units
3. Reliability of scale scores
4. Internationality of scale units
5. Comparability of scores on different scales
6. Ease of preparing duplicate scales

T scales and P.E. scales may be compared to physical measurements under the following categories:

Thing to Be Measured	Instrument	Attribute Measured	Unit	Zero Point
Man's height	Yardstick	Height	Foot and inch	Natural
Difficulty (arithmetic)	Woody Arithmetic Scale	Speed	Minute	Natural
Difficulty (reading)	Thorndike-McCall Reading Scale	Quality	S.D. value	Statistical
Handwriting	Starch Scale	Quality	P.E. value	Statistical

It will be observed that P.E. and S.D. scales are not scientifically precise. They possess a high degree of dependability, but their validity depends upon many assumptions that are predictable on purely statistical grounds, and statistics favor the group as against the individual. The educational scales are much more reliable for group characteristics than for individual attributes. The pedagogical fate of John Brown cannot be decided with the degree of finality that attends the measuring of a board or the temperature of a room. The method pursued in scale construction is undeniably scientific, but the results are approximative and only statistically exact.

3. Validity and Reliability of Tests. The test having been organized for content and instruction and scaled through the fixation of the zero point, the establishment of difficulty units, and the equalization of intervals

separating them, all test makers proceed to ascertain the validity and reliability of the measuring instrument. By validity is meant how truthfully the test measures the trait for which it has been created. Reliability expresses the degree of constancy obtaining between repetitions of the same test.

(*a*) *Validity.* Validity is predicated when a high correlation is obtainable between the test under advisement and an accepted criterion. Intelligence tests are judged for discriminative capacity by a comparison with the results obtained from the application of the Stanford Revision to the same group, or against the pooled results of other accepted scales. Myers' Mental Measure bases its validity upon such foundation. The test is also correlated with pooled teachers' judgments, mean scholastic proficiency, or both. The degree to which the distribution of the scores obtained corresponds to the normal distribution is also accepted as corroborative evidence. So is the progression yielded from grade to grade and from age to age, the assumption being that normal distribution and progression would conform to the sequences inherent in all mental functions or products. Negative proof is furnished by the relative absence of subjectivity in giving the test and in scoring it.

(*b*) *Reliability.* The test maker must also seek to ascertain the magnitude of the variable errors present in the scores when a test is repeated or another form of it is given. The causes that produce variability are

many. Some of these are the effect of practice or familiarity, lack of fidelity to instructions, myriad changes in the subject due to chance associations, poor health, antagonistic attitudes, and improvement in scholarship chargeable to the infinitesimal daily accretions. The latter may be neutralized by the use of statistical measures. One is self-correlation. The same test or a duplicate test is given after a suitable interval. If the correlation between the two sets of scores is high — ranging from .55 to .9 — reliability is postulated. Reliability may be increased by lengthening the test or by repeating it, in duplicate form, a set number of times. How much longer should it be made or, in lieu of lengthening, how many times should it be repeated? Spearman furnished a formula for the former.

$$N = \frac{rx\ (rl - l)}{rl\ (rx - l)}$$

N = Number of times to be lengthened
rl = Present correlation between two forms
rx = Coefficient of correlation between augmented form and present correlation

Suppose that the present correlation is .72 and the desired correlation .95; then

$$N = \frac{.95\ (.72 - 1.0)}{.72\ (.95 - 1.0)} = \frac{.95\ (-.28)}{.72\ (-.05)} = \frac{.266}{.036} = 7.4$$

This tells us that the mean of 8 (7.4) forms of the test would correlate with the mean of 8 (7.4) other equivalent forms to the extent of .95.

4. Standardization. A test is said to be standardized when norms or standards have been established. Standardization is accomplished when, the test having been stabilized, a large unselected group is examined. The results obtained are placed in an array and mean scores are obtained for grades and ages. The magnitude of the group to be tested has not been decided. McCall advises " averaging the scores of varying numbers of cases " to note " the resulting fluctuations in the averages. When the addition of, say, 100 new cases does not materially alter the previously determined norm, the norm has stabilized."

5. Summary. Test building is an arduous task, the more onerous because the material out of which it is fashioned, the products tested, and the entities measured, are subtly variable and inconstant. Choosing representative questions, problems, or exercises requires wide knowledge and keen insight. The educational products to be measured involve more than one variable. These are hopelessly intertwined, so that test makers are compelled to make serious compromises with the law of a single variable, which enjoins statisticians to measure one component power or function at a time. But this is not all. Test makers are hard put to it to construct scientific measures when handicapped by the lack of absolute zero points and by the necessity of resorting to purely mathematical units of measure. Add to this the day-to-day inconstancy of children's minds and the influence of emo-

tional and physical disrupting fluctuations, and it becomes evident that the measurement of achievement cannot be embarked upon with less than scholarly attainments. The success of the undertaking cannot compare with that attending the construction of a physical standard, owing to the forced departures from scientific accuracy. The test or scale is not a yardstick; and while it does not achieve accuracy of measurement, particularly when applied to individuals as against groups, it constitutes a dependable measure if used with the following reservations:

1. A child's pedagogical status should not be certified on results obtained from one test; we have seen that the coefficient of reliability of the best test is but .9.
2. Great latitude should be taken with respect to the significance of norms; even the Stanford Revision is fallible as respects the individual.
3. That quality test should be given preference which substitutes P.E. or S.D. values for judges' estimates as to the progression of difficulty evidence; the best of judges suffer from subjective disabilities.
4. That test should be chosen which has been standardized on an adequate sampling; McCall's instruction with respect to the magnitude of the group is unequivocal.
5. Pupils should be retested at suitable intervals;

constant errors due to psychological causes should be neutralized.

6. The presumption should lie in favor of the pupil's ability; no test yet devised is infallible for the individual.

(2) *Traditional versus Standardized Tests*

The dispraise bestowed upon the classroom examination is based upon its inherent weaknesses. The teacher or supervisor who formulates examinations has depended upon his or her judgment as to the kind of question to be incorporated. Uniformity is notoriously lacking; and judgment is a highly flexible attribute. Again, vagaries and divergence of judgment permeate teachers' attitudes to other fundamentals of measuring performance, as testing for specific things, testing under uniform conditions, and scoring with some degree of uniformity. There is wide disagreement as to the worth-while elements of a subject. Directions to pupils vary from mere exhortation to do their best to highly complex instructions teeming with ambiguities. Scoring methods are notoriously lax; the same answer has been known to receive ratings of 50% to 90%; and it is not an uncommon occurrence to find pupils graded in one subject poor by one teacher and bright by another.

On the credit side of the ledger may be placed the following advantages of homemade tests:

1. No extra expenses are involved.
2. Teachers are saved the energy and time of mastering the elaborate technique of many standardized tests.
3. They are not compelled to strive after a degree of efficiency imposed by wide-ranging norms of national scope; the teacher's primary concern is how well John Smith of Boston is doing as compared with other Bostonians and not with New Yorkers or, worse still, with a sectionally typical child.
4. Teachers can test the topic actually taught in the classroom. They are not forced to resort to the indirection of testing such proportions of it as the test maker had seen fit to incorporate, if indeed it finds a place at all.

Standardized tests are becoming increasingly free from the more serious defects inherent in the traditional tests. The personal equation does not weigh so heavily in the balance. The pupil's performance is measured with praiseworthy objectivity; an attitude of detachment, if not of impersonality, pervades the more recent tests. The tests, as instruments of measurement, are growing scientific in structure, though rigid accuracy eludes the most persistent of the test makers. There is a tendency to formulate both national and local norms, urban and rural standards. In a word, standardized tests are more dependable and, within the limits

aforementioned, are far superior to the traditional tests.

1. The New Type of Traditional Examination. There is a movement on foot to modernize the traditional test by making use of the technique popularized by the standardized tests. It does not aim to create a scientific instrument, but to reduce the drudgery, for both teachers and pupils, incident to the essay type of examination and to instill into them a modicum of objectivity. The new tests provide more explicit directions and simplify pupils' responses. Six forms are used, singly or in combination: (1) completion, (2) multiple choice (five possible responses), (3) multiple choice (three possible responses), (4) multiple choice (two possible responses), (5) true-false, (6) matching.

G. M. Ruch, who made an exhaustive study of the new-style examination, has published his results in *The Improvement of the Written Examination.** He finds that:

1. "The reliability of the objective types per unit of working time is very much greater than for the traditional essay type of examination, reliabilities of .80 to .90 being possible for twenty-minute tests in comparison with .40 to .70 for thirty- to sixty-minute written examinations."

2. "It is probably better to increase the length of objective examinations somewhat in order to eliminate

* Scott, Foresman, 1925.

chance successes rather than to give short examinations and depend upon statistical corrections."

3. "Guessing is probably relatively infrequent in examinations of the proper degree of difficulty. The proper degree of difficulty can be defined as that of allowing the average score to be approximately one-half the maximum possible score. No pupil should earn either a perfect score or a zero score. If he does, he cannot be said to have been measured."

4. The short-answer form of examination is more reliable than the traditional examination, the reliability coefficients for examinations of 100 items being:

Recall (completion)	.90
Multiple choice (5)	.89
Multiple choice (3)	.75
Multiple choice (2)	.85
True-false	.71

It is not to be assumed that these examinations are meant to duplicate the standardized tests. They are now being used as supplementary measures of achievement, filling the hiatus in the latter. They are designed to serve as a check upon the efficiency of the teacher's methods and the pupils' periodic attainments. Standardized tests serve other purposes.

2. The Use of Standardized Tests of Achievement. (*a*) *Supervision.* Achievement tests are used for supervision, for diagnosis, and for guidance. The supervisory uses are many; the administrative official may

desire to find out how his school compares with others, or how any grade of one class compares with another. The findings are tabulated and improvements are suggested; or perhaps teachers are rated, though sound pedagogical practice demands that other factors such as the mental status of pupils, methods employed, and the energy displayed by the teacher be included. Again, many schools promote pupils on the basis of such tests plus traditional examinations.

(*b*) *Diagnosis.* Many educational tests are designedly diagnostic; they purpose exploring the shortcomings of pupils within a narrow field of learning with a view to administering pedagogical remedies. It would seem that supervisors are losing sight of the fact that we cannot prescribe remedies unless and until we know the causes behind the symptoms. Especially is this true in the field of education, where the tracing of causes is a tedious and devious task. We shall examine two illustrative examples: diagnosis in arithmetic and in reading.

Suppose that a teacher tests a fourth-grade class with the Woody Arithmetic Scales, Series B, and finds that pupil A has fallen below the grade standards in all four operations. What procedure is indicated? Actually nothing beyond attesting to the inefficiency of the pupil; the diagnosis has merely a pedagogical significance, it simply shows that the instruction is not taking effect. Can the teacher formulate remedial measures? Unless remedial teaching is made synony-

mous with reteaching, no adequate program can be made. In order to ascertain the nature of the pedagogical malady revealed by the tests, the teacher must use the results as a starting point from which to traverse the whole psychological labyrinth beyond. The route to be covered includes the following paths:

1. Reorientation toward the psychology of learning arithmetic. This includes general intelligence, sensory acuity, long-distance memory, and formation of specific habits.

2. Survey of the pupil's mathematical disabilities. These may be traced to a low I.Q., poor hearing or eyesight or both, poor memory, ignorance of tables and other "bonds," emotional instability, presence of antagonistic attitudes toward a teacher or other pupils, tardy entrance in school, prolonged absences, or poor motivation.

3. Exhaustive investigation of the pupil's working habits in arithmetic. It sometimes happens that the pupil is carrying over infantile ways of counting, that he is addicted to computing by aid of "crutches," such as fingers, dots, etc., or that he had misunderstood the fundamentals from the very beginning of his school career, involving himself in serious mathematical misconceptions.

Let us consider a less comprehensive case. A child is asked to subtract $12\frac{5}{8}$ from 27 and gets $15\frac{5}{8}$ for the answer. Is the task of diagnosis completed by asserting that the pupil is poor in this phase of subtraction

and then proceeding to teach it again? Obviously not. For aside from the state of the pupil's intelligence and emotional stability, there exist several explanations for the lapse, each one demanding special treatment. Nor must we overlook the possibility that this failure may be due to entirely fortuitous causes — inattention, hurry, or indifference.

Diagnosis in reading is beset with similar pitfalls. A pupil who fails in the Gray Oral Reading Scale must submit to it for further investigation:

1. The level of intelligence:
2. Condition of his sensory, motor, and perceptual powers. Sensory elements include condition of the eyes — whether there is myopia, astigmatism, muscle weakness, or anomalies of the retina — and condition of the hearing. The motor elements are articulation, pronunciation, and eye movements. Perceptual elements involve the habits of responding to parts and to groups of words.
3. The emotionality of the pupil — whether attention is mobile enough, whether he possesses persistence and ability to concentrate.

The educational tests are not the self-diagnostic. They simply reveal a condition which must be traced to its ultimate causes before real diagnosis may eventuate. The standardized tests are more useful in diagnosis than the traditional tests because the prob-

lems and exercises that form the content of the former are more truly representative of fundamental facts and processes. But diagnosis can be made without resorting to them, so that the claim of possessing diagnostic properties cannot be substantiated.

A diagnostic chart in arithmetic. The Diagnostic Chart for Individual Difficulties in Fundamental Processes in Arithmetic* fulfills the major requirements of a good diagnostic test. The authors, G. T. Buswell and Lenore John, have studied critically the habits of children in addition, subtraction, multiplication, and division through the use of a Pupil's Work Sheet designed for individual testing of the quality of work in the fundamental processes. The pupil is required to work examples under the immediate supervision of the examiner, who admonishes him " to do all his thinking aloud." The examiner is thus enabled to add introspection to the traditional testing technique and to reveal the psychological basis of errors.

The Pupil's Work Sheet provides a graded series of examples in each of the four operations, arranged and tested out on five hundred school children. The gradation is from very simple to an upper segment of very difficult examples. The Diagnostic Chart, which lists faulty habits of work in order of frequency, is based upon an analytic study of the prevalent errors found in the work of two hundred fifty school children. It reproduces the examples of the Pupil's Work Sheet

* Published by the Public School Publishing Co., Bloomington, Ill.

and provides ample space for examiner's notations on each example.

Faulty habits are indicated for each of the fundamental operations. Thus under "Addition," 28 habits are named, ranging from "errors in combinations through error in writing answer," to "carrying when there was nothing to carry." The highest number of specific errors, 35, is listed for division. In this fundamental process a pupil's failure may be chargeable to such diverse faults as "errors in multiplication," "omitted final remainder," "used dividend or divisor as quotient," or "used second digit of divisor to find quotient."

To bring unity out of diversity, it may be well to group these habits under distinct categories. We find that many "habits" are condemned because they are uneconomical; some habits are caused by unfamiliarity with arithmetical facts (ignorance of combinations and tables); others are emotionally conditioned; and a few are due to inattention.

Examples of Uneconomical Procedure

(*A*) Adding Carried Number Irregularly

Example

879
266
498
167
137
1727

The pupil said, "8 and 9 is 17, and 6 is 23, and 7 is 30, and 7 is 37; 3 and 6 is 9, and 9 is 18, and 6 is 24, and 7 is 31 and 1 (carried) is 32; 8 and 1 (carried) is 9, and 2 is 11, and 4 is 15, and 2 are 17."

(*B*) Split Numbers

EXAMPLE

66
98
68
49
281

Here the pupil worked as follows, " 9 and 8 are 17, and 8 more — 17 and 4 are 21, and 2 are 23, and 2 more are 25, 3 more is 28, and 3 more is 31 ; 4 and 6 are 10, 10 and 9 are 19, and 6 are — 19 and 3 are 22, and 3 more 25, and 3 to carry, 28."

(*C*) Writing Rows of Zeros

EXAMPLE

5602
100
0000
0000
5602
560200

" 0 times 2 equals 0000."

(*D*) Counting to Get Quotient

EXAMPLE

8)56036

" 8 in 56 goes — 8, 16, 24, 32, 40, 48, 56 — 7 times."

EXAMPLES OF IGNORANCE OF ARITHMETICAL FACTS

(*A*) Errors in Combinations

EXAMPLE

132428
−38679
83749

In last combination, " 12 take away 3 is 8."

(*B*) Failing to Borrow

EXAMPLE

$$\begin{array}{r} 42 \\ 36 \\ \hline 10 \end{array}$$

"6 from 2 is nothing; 3 from 4 is 1."

(*C*) Using Wrong Operation

EXAMPLE

$2\overline{)46}$ "2 times 2 is 4; 2 times 3 is 6."

(*D*) Deriving Unknown Combination from Another

EXAMPLE

$$\begin{array}{r} 43 \\ 8 \\ \hline 344 \end{array}$$

"10 times 4 is 40, less 4 is 36, less 4 is 32."

ERRORS DUE TO EMOTIONAL BIAS

One pupil skipped around when adding a column of figures, explaining that he "did not like to add and therefore he always added the most difficult numbers first, in order to get through with them." There is much evidence that points to the adverse effect of emotional bias on school work. Facts that have been mastered to the point of overlearning are suddenly effaced; recall is retarded or blocked owing to the inhibiting influence of an associated emotional state. The hostile attitude of "I hate this subject" or "I hate this teacher" sometimes unsettles the most firmly implanted knowledge.

ERRORS DUE TO INATTENTION

(*A*) Ignoring a Digit

EXAMPLE

498
52
476

"2 from 8 is 6; 2 from 9 is 7; 0 from 4 is 4."

The 5 was ignored.

(*B*) Neglecting to Use Remainder within Problem

EXAMPLE

2)964(432

"2 into 9 goes 4 times; 2 into 6 goes 3 times; 2 into 4 goes 2 times."

(*C*) Securing Right Answer but Using Wrong One

EXAMPLE

9)41670(4407

"9 into 41 — 4 times and 5 over; 9 into 36, 4."

Diagnosis as here advocated by the authors leads the teacher deeper into the recesses of the child's mind than any other standardized test known to us. Thinking aloud by pupils is highly revelatory of hidden causes. It furnishes definite objectives for which to plan remedial teaching. But it does not sufficiently clarify the psychological factors behind the errors. The teacher can never feel certain where ignorance ends and emotional bias begins, for the latter may operate without the pupil suspecting its existence. Again, the pupil may be acutely conscious of affective disturbances and still feel reluctant to disclose them even to a

sympathetic examiner. Nor is there a distinct demarcating line between ignorance and inattention, or emotional bias and inattention. Besides, what warranty is there that the causative factors are at work permanently or sporadically? Isn't it possible that retesting the pupil may fail to substantiate the diagnosis?

Tests like the Diagnostic Chart are more dependable than tests without introspective confirmation, but they do not relieve us of the duty to establish, by repeated testing, such traits of personality as (1) long-distance memory, (2) the I.Q. intelligence level, and (3) emotional and temperamental coloring as disclosed by observational records and complex-indicators.

(*c*) *Guidance.* Guidance is either educational or vocational. Educational guidance means the direction of a pupil's scholastic career from the degree of predictibility evident in present performance. Vocational guidance refers specifically to vocational counseling as suggested by scholastic trends, revealed through tests. Guidance calls for a prophetic faculty which enables the practitioner to foretell future performance. Opinions as to the prognostic value of scholastic efficiency vary widely. Hollingworth is inclined to sound a note of optimism. "Those who are destined to achieve distinction and success begin to do so at an early age. . . . In spite of comments to the contrary, the school curriculum would seem to constitute a useful test in prognosticating at least the most probable qualities of the individual's later work." On the other hand,

others cast doubts upon the early florescence of a career. There is many a slip between promise and success. Absorbing interests may supervene and wreck many a resolution. The absence of desirable personality traits of a non-intellectual character may rob a pupil of the success to which his scholarship entitles him, so that the one who plays the rôle of vocational prophet must inure himself to failure.

There is one aspect of vocational guidance that deserves serious consideration. This is the emphasis being placed by executives on personality in its non-intellectual phases. Prospective employers are disposed to discount scholarship or even specific knowledge and to stress character. Given intelligence above moronity, the applicant must satisfy the employment superintendent as to his moral qualities. This has been the experience of Jesse B. Davis. In his book *Vocational and Moral Guidance*, he bears witness to this tendency. " When I asked the employer whether he wished the applicant to know certain things, he invariably replied: ' We will teach him the things we wish him to know. What I want you to do is to help me find the right kind of a fellow.' "

On the whole, it would seem rash to prognosticate from educational tests. This does not mean that guidance should be abandoned. But this should be done from long perspective and not without balancing all personality traits. The rule should be to guide the total individual.

SUMMARY OF REPRESENTATIVE ACHIEVEMENT TESTS

SUBJECT	NAME OF TEST OR SCALE	BRIEF DESCRIPTION	GRADES FOR WHICH DESIGNED	TYPE OF NORM	PUBLISHER
Arithmetic	1. Buckingham Scale for Problems in Arithmetic	Reasoning in terms of difficulties; value based on the P.E. location	3 to 8	Grade	Public School Publishing Co., Bloomington, Illinois
	2. Courtis Standard Practice Tests in Arithmetic	Practice in the form operations to bring up level of efficiency; form pads and self-corrective cards; 44 lessons in all; Student's Daily Record	4 to 8	Grade	S. A. Courtis, 1807 E. Grand Blvd., Detroit, Michigan
	3. Courtis Standard Research Test in Arithmetic, Series B	Four consecutive pages, each devoted to a fundamental operation	4 to 8	Grade	
	4. Monroe Diagnostic Tests in Arithmetic	Twenty-one tests in four parts; all operations; test both rate and accuracy — rate, number attempted; accuracy, per cent of the examples correct	4 to 8	Grade	Public School Publishing Co., Bloomington, Illinois
	5. Monroe Standardized Reasoning Test in Arithmetic	Test 1, grades 4 to 5; Test II, grades 6 to 7; Test III, grade 8. Problems scored for correct principle (C); no time limit; when time limit is used, rate sum of P values done correctly in 10 minutes	4 to 8	Grade	

	6. Stevenson Arithmetic Reading Test (Problem Analysis)	Problem analyzed under A, B, C, and D. (A) Which of the following facts are given in this problem? (B) Which of the following things are you asked to find out? (C) Which of the following is the most reasonable answer? (D) Which process should be used? Pupil checks one of four suggested answers	4 to 9	Grade	Public School Publishing Co., Bloomington, Illinois
	7. Woody Arithmetic Scale	Two Series, A and B, in four operations; Series A more reliable and longer than Series B	3 to 8	Grade	Teachers College, Columbia University, New York City
	8. Woody-McCall Mixed Fundamentals	Two forms; fundamentals; both forms complementary	3 to 8	T scores	Public School Publishing Co., Bloomington, Illinois
Geography	1. Courtis Standard Supervisory Test in Geography	Consists of 4 tests — (*a*) oceans and continents, (*b*) states, (*c*) cities, and (*d*) countries; method of scoring based on a 1000-point scale; all tests given in space of 6 minutes	4 to 7A	Grade	S. A. Courtis, 1807 E. Grand Blvd., Detroit, Michigan

SUMMARY OF REPRESENTATIVE ACHIEVEMENT TESTS (*Continued*)

SUBJECT	NAME OF TEST OR SCALE	BRIEF DESCRIPTION	GRADES FOR WHICH DESIGNED	TYPE OF NORM	PUBLISHER
	2. Gregory-Spencer Geography Tests (Forms A, B, and C)	Eight tests covering (*a*) trade routes and their products, (*b*) miscellaneous geography, (*c*) causal geography, (*d*) places and descriptive geography, (*e*) political and place geography; consume 4 minutes; contain 111 items; scoring by two-card stencil	6, 7, and 8	Grade	Bureau of Educational Research, University of Oregon, Eugene, Oregon
	3. Hahn-Lackey Scale	Takes the form of the Ayres Spelling Scale; furnishes teacher with classified list of questions; merely a question paper; scoring method not satisfactory	4 to 8	Grade	Public School Publishing Co., Bloomington, Illinois
Handwriting	1. Gettysburg Edition of the Ayres Scale	Specimens from value 20 to 90 in steps of ten; based on legibility in three varieties of slant; rate measured by number of letters written in two minutes; copy same for all specimens — part of Gettysburg address; scales posted in classroom to encourage auto-criticism	All grades		Russell Sage Foundation, New York City

2. Courtis Standard Practice Tests in Handwriting (to be used with the Ayres Scale): (1) Student's Daily Lesson Book (one for each pupil); (2) Student's Daily Record; and (3) Teacher's Manual	A self-scoring instrument; pupils practice on single words, sentences, and short paragraphs; may be used with Ayres Scale or a scale similar to it; the latter is made up of single letters and words instead of a set paragraph; there are four such scales, each ranging in value from 30 to 90; two research tests are included in which pupils copy a paragraph and compare with Ayres Scale	3 to 8	Individual	The World Book Co., Yonkers, New York
3. Freeman Diagnostic Chart	Five different scales, one each for (*a*) letter formation, (*b*) quality of line, (*c*) spacing, (*d*) uniformity of slant, (*e*) uniformity of alignment	All grades	Individual diagnostic scale	Houghton, Mifflin Co., Boston, Massachusetts
4. Thorndike Handwriting Scale	Measures (*a*) legibility, (*b*) beauty, (*c*) general merit; samples assigned numerical values from 4 to 18, with one or more specimens furnished for each degree of quality	2 to 8	Standards shown in the form of a graph combining rate and quality	Teachers College, Columbia University, New York City

SUMMARY OF REPRESENTATIVE ACHIEVEMENT TESTS (*Continued*)

SUBJECT	NAME OF TEST OR SCALE	BRIEF DESCRIPTION	GRADES FOR WHICH DESIGNED	TYPE OF NORM	PUBLISHER
History	1. Harlan Test of Information in American History	Ten exercises based on study of contents of textbooks on American history; no alternative forms; may be done in 25 minutes	7 to 8	Grade	Public School Publishing Co., Bloomington, Illinois
	2. Van Wagenen History Scales, Extensive Manual entitled "Historical Information and Judgment of Elementary School Pupils"	Three scales, information, thought, and character judgment; scales test (*a*) ability to retain facts, (*b*) ability to draw inferences, and (*c*) ability to evaluate character traits in a historical setting; duplicate scales; Character Judgment Scale L supplements Character Judgment Scales A and B; difficult to score because of partial credits	4 to 8	Grade. Different norms for boys and girls	Teachers College, Columbia University, New York City
Language	1. Briggs English Form Tests, Alpha and Beta	Elements of English composition — initial capital, the terminal period, the terminal interrogation point, the capital for a proper noun or adjective, the comma before *but*, coördinating members of a compound sentence; each test	7 to 9	Grade	Teachers College, Columbia University, New York City

	composed of the 7 elements, organized into a cycle, is used; time — 20 minutes; scoring by tissue paper stencils			
2. Charters Diagnostic Language and Grammar Test — Two Tests: (1) The Diagnostic Language Test (Grades 3 to 8); (2) Diagnotics Language and Grammar Test (Grades 7 and 8)	Each test composed of two kinds of sentences, some correct and some incorrect; pupils of lower grades indicate correct samples by making a cross on line below and incorrect samples by writing correct words on line; those in upper grades also state reasons; no time limit	3 to 8	Grade	Public School Publishing Co., Bloomington, Illinois
3. Hudelson Scale	Fifteen specimens; grade norms rise in value .6 of a unit; pupil's work compared with samples of scale; measures general merit	4 to 6	Grade	The World Book Co., Yonkers, New York
4. Lewis Special-Type Scales	Five scales — 2 are business letter scales, 2 are friendly letter scales, 1 is narrative compositions written on topic "One of My Most Interesting Experiences"	5 to 12	Grade	The World Book Co., Yonkers, New York

SUMMARY OF REPRESENTATIVE ACHIEVEMENT TESTS (*Continued*)

SUBJECT	NAME OF TEST OR SCALE	BRIEF DESCRIPTION	GRADES FOR WHICH DESIGNED	TYPE OF NORM	PUBLISHER
	5. Nassau County Supplement	Hillegas Scale improved by Thorndike (eliminated artificial samples, substituted school work, samples lengthened, and number increased) and Trabue; measures general merit; 9 samples — 7 written on one subject; difficult to use by untrained	Upper grades and lower high school	Individual norms. Score is value of sample to which pupil's work corresponds	Public School Publishing Co., Bloomington, Illinois
	6. Trabue Completion Test (Scales B, C, E, and T are equivalent. Scales Alpha and Beta are to be used individually.)	Sentences from which one or more words are omitted; sentences arranged in order of difficulty; scoring done by means of a key	2 to 8	Grade	Public School Publishing Co., Bloomington, Illinois
Reading	1. Burgess Scale for Measuring Ability in Silent Reading. Picture Supplement Scales I, II, III, and IV (alternative forms)	Dominant variable is rate, therefore all units of same difficulty; pupil shows understanding of pictures by following directions, which are to make marks with a pencil; abundantly supplied with statistical aids; can be given in 5 minutes; effect of practice is to increase score markedly	3 to 8	Grade and percentile	Russell Sage Foundation, New York City

2. Gray Oral Reading Test	Individual test; consists of 12 paragraphs; scores six types of errors: (*a*) gross errors, (*b*) minor errors, (*c*) omissions, (*d*) substitutions, (*e*) insertions, (*f*) repetitions; rate and power test combined, *i.e.*, second-grade child must read paragraph 1 in 40 seconds with less than five errors, credit being increased with lessening of time, decreasing of errors, or both	1 to 8	Individual	Public School Publishing Co., Bloomington, Illinois
3. Haggerty Reading Examination Sigma 1	Two tests plus 1 fore-exercise; composed in the terminology of second-grade children; not easily scored; alternative forms	1 to 3	Grade	The World Book Co., Yonkers, New York
4. Monroe Standardized Silent Reading Tests (Revised)	Test I for grades 3, 4, 5; Test II — 6, 7, 8; three forms for each test; tests difficulty or comprehension value; sum of comprehension values constitutes the score; each exercise also has a rate value; sum of rate values child tries in 4 minutes yields a rate score (words per minute)	3 to 8	Grade	Public School Publishing Co., Bloomington, Illinois

SUMMARY OF REPRESENTATIVE ACHIEVEMENT TESTS (*Concluded*)

SUBJECT	NAME OF TEST OR SCALE	BRIEF DESCRIPTION	GRADES FOR WHICH DESIGNED	TYPE OF NORM	PUBLISHER
	5. Pressey First-Grade Reading Vocabulary Test	Pupil required to pick out a *real* word in a number of real and nonsense words; to be located with assistance of teacher; child draws line around *real* words; preceded by fore-exercises	1A and 1B	Grade	O. R. Chambers, 215 E. 3d St., Bloomington, Indiana
Spelling	1. Ayres Measuring Scale for Ability in Spelling	Efficiency in spelling based on 1000 calibrated words	3 to 8	Grade	Russell Sage Foundation, New York City
	2. Monroe Timed Sentence Test	Spelling efficiency in which words from Ayres list are embodied	3 to 8	Grade	Public School Publishing Co., Bloomington, Illinois

Exercises for Self-Testing and Investigation

I. 1. Pressey Classification tests
   (*a*) arithmetical ability.
   (*b*) historical judgment.
   (*c*) general ability.

2. Pressey Diagnostic Tests in English Composition are
   (*a*) for grades 3–4.
   (*b*) for grades 9–12.
   (*c*) for grades 7–12.

3. The Hudelson Typical Composition Ability Scale tests
   (*a*) narration.
   (*b*) paragraph structure.
   (*c*) general merit.

4. Willing Scale for Measuring Written Composition measures
   (*a*) general merit.
   (*b*) form value and story value.
   (*c*) power of description.

5. Clark Letter Writing Test applies to
   (*a*) friendly letters.
   (*b*) applications for positions.
   (*c*) social and business letters.

6. Monroe Standardized Silent Reading Tests are for
   (*a*) comprehension.
   (*b*) speed.
   (*c*) rate and comprehension.

7. Lister-Myers Penmanship Scale tests
   (*a*) legibility, beauty, ease.
   (*b*) uniformity of slant, uniformity of alignment, quality of line, letter formation, spacing.
   (*c*) general merit.

8. Hillegas Scale contains
   (*a*) four compositions.
   (*b*) fifteen compositions.
   (*c*) ten compositions.

9. Van Wagenen American History Scales
   (*a*) consist of one composite test.
   (*b*) test information, thought, character, judgment.
   (*c*) test information, thought, historical sense.

10. Stevenson Arithmetical Reading Test is a test of
   (*a*) silent reading based on arithmetic.
   (*b*) problem analysis.
   (*c*) arithmetical terms.

II. The three numbered lists subjoined represent authors, names of tests they have devised, and characterizing sentences or phrases descriptive of them. These do not occur seriatim. They are to be arranged in tabular form under the headings A, B, and C. For instance, the Gray Oral Reading Scale, a test in oral reading for all grades, the respective numbers of which in divisions A, B, and C are 9, 1, 7, should appear as:

A	B	C
9	1	7

A

1. Courtis
2. Monroe
3. Buckingham
4. Halley
5. Burgess
6. Thorndike-McCall
7. Ayres
8. Starch
9. Gray
10. Charters
11. Hillegas

B

1. Oral Reading Test
2. Sentence Vocabulary Scale
3. Standard Research Tests, Series B
4. Spelling Scale
5. Handwriting Scale
6. Diagnostic Tests in Arithmetic
7. Composition Scale
8. Picture Supplement Scale
9. Scale for Problems in Arithmetic
10. Reading Scale
11. Diagnostic Language Tests

C

1. Includes a fairly complete list of frequent types of language errors.
2. Provision is made for converting the point score into a T score.
3. Known as the Gettysburg Edition.
4. The words were chosen by taking the first defined word on the even-numbered pages in Webster's New International Dictionary (1910).
5. Composed of four tests in each of the fundamental operations in arithmetic and are both diagnostic and general.
6. Problems located on basis of P.E. values.
7. A test in oral reading for all grades.
8. A series of twenty-one tests confined to fractions and decimals.
9. It is typified by "An orange is a . . . dress . . . animal . . . fruit . . . hornet."
10. It consists of twenty exercises in making directed supplementary drawings in connection with a picture.
11. The scale values of this measuring instrument range from zero to 9.3.

Bibliography

Ashbaugh, E. J. — *Handwriting of Iowa School Children*, University of Iowa, 1916

—— "The Measurement of Language," *Journal of Educational Research*, June, 1921

Ayres, L. P. — *A Measuring Scale for Ability in Spelling*, Russell Sage Foundation, 1915

Buckingham, B. R. — *Spelling Ability, Its Measurement and Distribution*, Teachers College, Columbia University, 1913

Burgess, M. A. — *Measurement of Silent Reading*, Russell Sage Foundation, 1921

Chapman, J. C., and Counts, G. S. — *Principles of Education*, Houghton Mifflin, 1924

Gray, C. T. — *Deficiencies in Reading Ability*, Heath, 1922

Gregory, C. A. — *Fundamentals of Educational Measurement*, Appleton, 1922

Griffith, G. L. — "Harlan's American History Test in the New Trier Township Schools," *School Review*, November, 1920

Hines, H. C. — *A Guide to Educational Measurements*, Houghton Mifflin, 1923

Hollingworth, L. S. — *The Psychology of Special Disability in Spelling*, Teachers College, Columbia University, 1918

—— *Special Talents and Defects*, Macmillan, 1923

Hudelson, E. — "English Composition; Its Aims, Methods, and Measurement," *Twenty-second Year Book*, Bloomington, Illinois, 1923

Klapper, P. — *Teaching of Reading*, Appleton, 1926

McCall, W. A. — *How to Measure in Education*, Macmillan, 1922

Meade, C. D. — *An Experiment in the Fundamentals*, World Book, 1921

Monroe, W. S., De Voss, J. C., and Kelly, F. J. — *Educational Tests and Measurements* (Revised), Houghton Mifflin, 1924

Otis, A. S. — "The Reliability of Spelling Scales," *School and Society*, October 28, November 4, 11, 18, 1916

Paulu, E. M. — *Diagnostic Testing and Remedial Teaching*, Heath, 1924

Rice, J. M. — "The Rice Spelling Tests," *The Forum*, March, August, 1897

Russell, C. — *Classroom Tests*, Ginn, 1926.

Sears, J. B. — "Measurement of Teaching Efficiency," *Journal of Educational Psychology*, September, 1921

Seashore, C. E. — *The Psychology of Musical Talent*, Silver Burdett, 1920

Stone, C. R. — *Silent and Oral Reading*, Houghton Mifflin, 1922

Suzzallo, H. — *The Teaching of Spelling*, Houghton Mifflin, 1913

Thorndike, E. L. — *The Psychology of Arithmetic*, Macmillan, 1922

Trabue, M. R. — *Measuring Results in Education*, American Book, 1924

—— "Supplementing the Hillegas Scale," *Teachers College Record*, January, 1917

Van Wagenen, M. J. — *Historical Information and Judgment of Elementary School Pupils*, Teachers College, Columbia University, 1919

Williams, J. H. — *Graphic Methods in Education*, Houghton Mifflin, 1924

Wilson, H. B., and Wilson, G. M. — *Motivation of School Work*, Houghton Mifflin, 1921

Woody, C. — *Measurement of Some Achievements in Arithmetic*, Teachers College, Columbia University, 1916

Zirkes, L. — "Diagnostic Measurement as a Basis for Procedure," *Elementary School Journal*, March, 1918

CHAPTER VII

TESTING NON-INTELLIGENCE TRAITS OF PERSONALITY

The average teacher or administrator senses but vaguely the multi-dimensional nature of personality. Current psychological literature, especially the unprecedented vogue of Freudianism and its variants, has increased their familiarity with emotional states (complexes), temperamental disturbances, and correlative bodily conditions. They assign priority, however, to intelligence traits and neglect or slur other traits because of the feeling (1) that they are the special province of psychiatry — an occult and recondite branch of psychology, (2) that psychiatrists, like the professor in the *Water Babies*, go exploring with a probe in search of a mysterious and elusive "hippopotamus major," (3) that the lack of mathematical formulation and foolproof standardization throws too many hazards in their path and redoubles their faith in tests "which any teacher who can understand directions can give," and (4) that the normal child is afflicted little, if at all, with emotional disturbances. Many psychologists have helped to sow these doubts by making preposterous claims for intelligence tests. Not that they ignore

what Haggerty calls " non-intelligence traits," but they give the impression that this one-dimensional measure is all-sufficing in predicting normality and potential scholastic competence; in cases where a discrepancy occurs between intelligence and performance they supplement the findings of intelligence tests by purely descriptive terms, as " emotional disturbance " or " lack of industry."

Fortunately for the measuring movement, psychiatry has invaded the province heretofore appropriated by the psychologists and mental testers. Their relations are fast becoming cordial. The rapprochement has been effected by the gradual modification of the disparate terminology that fortified each in their adherence to their favorite instruments of testing, and by extending the meaning of " personality " to include all aspects of child nature.

The swelling rank of testers has been augmented by still other accessions. The sociologist has made no less noteworthy contributions. He maintains, and rightly so, that since the child develops in a social milieu, it is impossible to envisage personality completely without taking note of factors making for sociality, comprehended under the term " togetherness." He insists on knowing whether the individual surveyed can work readily with his fellows, whether his coöperative behavior is marred or enhanced by docility or aggressiveness, for intellectual attainments may be nullified by the inability to " float in society."

Nor is this the only quarrel the sociologist has with the mental tester. Society, he asserts, molds each individual differently and builds around him a universe through the intimate interplay of innate and environmental forces. Since the quality of each is different for each individual, it follows that for each individual there exists a distinct universe with characteristics of its own. As Professor Green acutely remarks: "If we have two individuals A and B, we have two universes to consider, *a* and *b*. If the mental efficiency of A in universe *a* is *x* and that of B in *b* is *y*, we cannot compare the efficiency of A and B until we know the relative complexity of the two universes."

The sociologist is ably seconded by the physiologist and anatomist. Both physiology and anatomy have yielded the information that the tempo of bodily growth varies considerably, maturation being retarded or accelerated in many individuals. The period of adolescence may be of brief or long duration. Many a child who tests "borderline" in mentality may be suffering from a corresponding retardation in bodily growth; his tempo of development may become accelerated at a later period to the point of normality. This group of retardates is matched by an equally large group of seeming precocities. They may test very high mentally and therefore be classifiable as "precocious." Yet both physiological and anatomical ages may be far above their mental ages: a ten-year-child who tests at twelve may have the physiological development of a

twelve-year-old. These considerations should give pause to the mental tester, particularly in his treatment of that type of child who matures at an early age and contracts the period of adolescence — a time of most delicate and complicated adjustments that cannot be hurried through. Ogden sums up their plight succinctly: "If they are 'bright' children, they may show startling signs of genius at an extraordinarily early age only to flatten out in adulthood and become quite commonplace personages. If they are 'dull' children they are still more serious problems to society, for they mature before they can know enough to get properly adjusted to the world." Cyril Burt offers similar evidence gleaned from the study of the young delinquent who outgrows his schoolmates.

A. The Educational Significance of Non-Intelligence Personality Traits

Intelligence is not a primary force. Back of intelligence is the mind, and behind the mind is the body. While intelligence is not purely derivative, it is motivated by traits that are external to it. These non-intelligence traits belong to the category of emotions, volitions, and bodily states. Intelligence apprehends and assimilates; emotions give tone and vitality to these functions; the will regulates their degree of intensity; and physical well-being places a limit to their continuance. Disturbances in emotional balance, infirmities of will, and diseases of the body are reflected

in the mind; and the extent of mental aberration ranges from mild psychopathy to virulent states of dementia. Intelligence flounders like a ship that does not answer the helm; no amount of mechanical improvement of ship or engine can be substituted for a broken rudder.

Educators have long been aware of this close intimacy. Children with high I.Q.'s prove educational laggards, and honor pupils cannot always boast of a flattering I.Q. There are pupils who could if they would, and many who would if they could, and still others in whom the will-to-do bridges the gap between intellectual poverty and scholastic excellence. Sometimes the explanation is close to the surface. Self-consciousness, laziness, and malnutrition account for many discrepancies between intellectual promise and pedagogical performance. Sometimes the cause may be due to ineptness. What sometimes demarcates the borderline case from the normal is the manner of employing the mental capital at the disposal of each; the abnormal child expends it recklessly, extravagantly, and out of all proportion to the results obtained. In this respect he is not unlike Washington Irving's athlete, who in a hurdle competition took a start of three miles. He had drained his vitality before reaching the point of maximum exertion.

At other times the cause may be too baffling for even trained observation. It may exist in the subliminal depths of subconsciousness — below the level of the conscious layer, below the level of the memory layer.

It is chargeable to the maladjustment between the changeless biological organism working itself out through instincts and emotions, to the changing conditions of civilized society. Continued maladjustment may issue in conflicts and repressions that may lead to dissociations of personality into a dualism of the Jekyll and Hyde variety, as each one alternates in periodic domination of behavior. Educators have not as yet discovered the healing potion that Stevenson provides for his famous character. Upon psychiatry and psychoanalysis devolves the task of removing the cause. The latter has perfected a " technique " which, when used by experts, has cured many a difficult " case " of discipline, and has restored to full mental vigor the depleted energies of many " nervous " children.

Unfortunately, the measurement of intelligence is in danger of being discredited by its devotees, whose zeal betrays them into psychological indiscretions. The seductiveness of objective evaluation, the subtle appeal of statistical figures, tend to obscure educational thought. Proficiency in school work is not always conjoined with a high intelligence quotient, for fine native ability may reside in an individual who lacks certain character traits which alone can transmute potential intelligence into actual intelligence. Then, too, the emotions may enhance or lower scholastic attainments. Perseverance, the capacity to cleave to a line of conduct, may compensate for indifferent mental endowment; industry enhances one's intel-

lectual gifts. Lack of emotional balance — extreme irritability and self-consciousness — may unfit a child for participation in class instruction. Indeed, the emotionally inept do not show up to good advantage in any but individual intelligence tests. Group tests do not allow the fine adjustments personal contact permits; truculence or vacillation passes undiscovered; the numbing effect of overexcitement, damming up for a time the stream of consciousness, can best be disclosed by cross-examination or introspection.

That intelligence and scholarship are not always interdependent is no longer disputed by educators. Those who, in the first blush of enthusiasm, extolled intemperately the infallibility of intelligence tests have revised their judgment. Their enthusiasm abated considerably when experience and experiment had demonstrated the insufficiency of data supplied by intelligence tests alone, for the learning process is carried on by the whole child, not merely by his brain — by the will and the emotions as well as by the intellect.

Notable pioneer work in the testing of character was undertaken by Lasurski, Pressey, and Downey. The contributions of each typify a class or " school " of tests.

Pressey made a painstaking investigation of the interrelations subsisting between intelligence, achievement, preparation, and " school attitude." He found that the degree of intelligence bears no causal relation to scholastic achievement. Not so with " school atti-

tude." The pupil whose volitional and emotional attitude toward the teacher and the school tasks insures a proper, effective tone succeeds best in scholarship.

While Pressey's contribution to educational prognosis is considerable, he has not freed his non-intelligence tests from the fluctuations of subjectivity. His rating blank does not lend itself to concrete estimation of character traits. The probable error in teacher judgment is very high; a coefficient or reliability of 40 indicates too wide an amplitude. Nevertheless, it possesses merit of a mathematical definiteness, for the qualities are given numerical values.

Lasurski devised a record for the observation of the subjective aspects of personality. Eleven traits are recognized: (1) attentiveness, (2) adaptability, (3) memory, (4) imagination, (5) thinking, (6) language, (7) general intellectual characteristics, (8) feeling, (9) general emotional characteristics, (10) volitional traits, (11) resolution.

While Lasurski does not demarcate intelligence from non-intelligence traits, he gives adequate space to the plotting of the latter. Other investigators have striven for more definiteness and concreteness through the detailed analysis of each trait. The Badanes Observational Record offers the teacher minute directions for the comprehensive survey of those subjective qualities determinative of scholastic proficiency. The method pursued is one of suggestive descriptions, condensed

into pithy phrases, organized into logical units, and designed to disclose countless facets of personality. It succeeds in probing deeply into the dim recesses of character and emotions; it unearths the hidden springs that feed the intellectual current.

The exact title of the Badanes Observational Record is "A Psychological Pedagogical Observational Record for the Selection of Gifted Children in the Elementary Schools." It may be used with equal competence in evaluating personality traits of normal and slow children. It seeks information on (1) general condition of senses and the nerves, (2) psychic attributes, (3) imagination, (4) ability to think, (5) emotions, (6) the will, (7) language, (8) manner of work, (9) relation to community and home, (10) interests and abilities, (11) adaptation to new demands, and (12) a full statement of the characteristics of the pupil. It should be borne in mind that the gifted seem to excel in all qualities, both physical and psychic, to an equal degree, so that from the presence of one may be inferred the possession of allied attributes. Not so with the normal and the slow. An observer of the latter groups can derive no corroborative aid from the high — or low — correlation of characteristics that obtains among the gifted. He must study each personality trait as an independent phenomenon, the existence or absence of which he cannot take for granted.

Excerpts * under the major head will further elucidate its technique:

General Conditions of the Senses and the Nerves

Aim of Observation	Opportunities for Observation	Decisive Characteristics or Attributes	Remarks
Nervous irritability	(1) The way he sits or stands	(1) Quiet or fidgety	
	(2) The way he impresses you at the end of the morning, in the afternoon, end of week, end of term, after a period of arithmetic or gymnastics	(2) Very easily fatigued, easily fatigued, or difficult to fatigue, physically and mentally	
	(3) His reaction during a thunderstorm, during extreme heat, or during change of weather	(3) Very little influenced, moderately influenced, considerably influenced	
	(4) His reaction before, during, and after examinations, excursions, and school festivals	(4) Highly excited or calm	
Attention			
(*a*) Is the attention aroused?	Involuntary attention of child	Easy, medium, or slow adjustment of attention	
(*b*) Does the pupil possess the ability to concentrate?	Effect of slight disturbance, such as noise by pupils or noise from street, effect of slight physical discomforts or slight pain	Unstable, fluctuating, or irregular attention	
(*c*) Is the pupil's attention easily distracted?		Fixative or fluctuating	

* Reproduced with the permission of the publisher, Saul Badanes.

GENERAL CONDITIONS OF THE SENSES AND THE NERVES (*Continued*)

AIM OF OBSERVATION	OPPORTUNITIES FOR OBSERVATION	DECISIVE CHARACTERISTICS OR ATTRIBUTES	REMARKS
Memory and learning			
(*a*) Does the pupil memorize in a mechanical manner, word for word (verbally), without paying attention to the meaning, or does he memorize according to the content and meaning, memorizing and retaining the essentials?	Mechanical memory shows itself when he recites what he has learned, making absurd errors and emphasizing the wrong point	Mechanical (verbal) or logical (judicious)	
(*b*) Is the pupil's memory developed equally well for names, words, numbers, memory gems, persons, colors, pictures, melodies, places, and abstract terms, or only for special subjects?	In all subjects	A one-sided or many-sided memory, good for names, words, numbers, memory gems, etc.	
Imagination			
Does the pupil possess a creative imagination?	Inventing new ways of playing, nicknaming, drawing, making melodies, telling stories	Highly creative, moderately creative, slightly original	
Ability to think			
(*a*) Does the pupil possess the ability to combine concepts and thoughts and find the true relation between them?	Solving riddles or puzzles, forming sentences from one or more given words, filling blanks in sentences, explaining and applying proverbs and figures of speech, solving problems independently	Strong, medium, or weak power of combining	

GENERAL CONDITIONS OF THE SENSES AND THE NERVES (*Continued*)

AIM OF OBSERVATION	OPPORTUNITIES FOR OBSERVATION	DECISIVE CHARACTERISTICS OR ATTRIBUTES	REMARKS
(*b*) Can the pupil work independently or does he need help?	Preparing home work without help, independent solving of problems, understanding of reading matter not taken up in the classroom, expressing original thought in composition	Able, slightly able, or not able to do independent intellectual work	
Emotions			
(*a*) Egoistic sentiments, especially the knowledge of his own worth and his own sense of humor	(1) The way he acts towards his teacher and classmates (2) The way he behaves when he or another pupil is praised or censured (3) When teased, whether he scolds or answers back (4) When praised, whether he is stimulated to greater effort		
(*b*) Altruistic sentiments, especially sympathy	His concern for the welfare of others (interest in friends and enemies)	Pity, sharing joys and sorrows with others, envy, readiness to help, regard for others, like or dislike of by others, number of friends	
(*c*) Intellectual and æsthetic emotions	(1) His enjoyment of instruction (2) When he sees pictures or beautiful landscapes, or hears beautiful music	Outspoken pleasure in acquiring knowledge Pleasure in seeing or hearing the beautiful	

GENERAL CONDITIONS OF THE SENSES AND THE NERVES (*Continued*)

AIM OF OBSERVATION	OPPORTUNITIES FOR OBSERVATION	DECISIVE CHARACTERISTICS OR ATTRIBUTES	REMARKS
(*d*) Moral sentiments	(1) In social life of school		
	(2) In subject matter that has ethical content	Lively moral judgment, sense of approval of good behavior	
	(3) His behavior toward his classmates, especially in reference to lying, dishonesty, and injustice	Outspoken sense of truth, keen sense of justice or lack of it	
The will			
(*a*) Is the pupil lively and active or indifferent and phlegmatic?	(1) Discharge of all claims or demands of school	Strong, medium, or weak; energetic or languid; slow, lukewarm, or quick	
	(2) Energy with which he carries out his purposes; in competitive games	Inactive, languid, or indolent	
	(3) His behavior when he is hindered, or in difficulties, when he is sick, or when he is disturbed by other people		
	(4) Readiness for mental and physical movements in rising, walking, running, etc.		
(*b*) Does the pupil persevere in difficulties or does he stop trying soon?	(1) In subjects where he accomplishes little	Persevering, consistent, orderly, thorough, unceasing effort; wavering, inconsistent, superficial, fickle, inconstant	
	(2) His attitude toward his own desires and interests		

General Conditions of the Senses and the Nerves (*Continued*)

Aim of Observation	Opportunity for Observation	Decisive Characteristics or Attributes	Remarks
	(3) Whether he begins many things without finishing them		
Special interests and talents			
(*a*) Has pupil special interests outside of school work (handiwork, nature work, gardening, music)?	By questioning pupils and parents	Objective, subjective, or both	
(*b*) What does he read?	School and home reading	With or without discrimination, once or more, yes or no	
(*c*) Does he read indiscriminately any book he gets hold of? Does he read them again? Does he dramatize in games and play the books he has read?	Same as above	Same as above	

The Observational Record, however, is too circumstantial in its observations and entails an enormous expenditure of time. The " pupil-load " that the exigence of modern school organization assigns to teachers is far too heavy to increase the burden. The German psychologists, who originated the observational record, have been forced to reduce its proportions. The educational authorities of Berlin content themselves with a brief account of the pupil's salient characteristics. Dr.

Müller published a series of excerpts in the April, 1922, number of *Praktische Psychologie*, from which we cull three samples:-

1. "W. H. displays a fine type of apprehension, good memory, and ability to reproduce lessons faithfully. He is well-behaved, takes keen pleasure in his work, and is persevering in the mastery of difficult tasks. He is unobtrusive, serious, and painstaking in his written exercises."

2. "Karl B. has always outranked his classmates. This pupil is very alert, possesses the power of rapid adjustment, has a thirst for knowledge, and is endowed with critical capacity. His written work is characterized by thoughtfulness and system. His attitude towards his schoolmates is all that is desired."

3. "Meta R. — *Scholastic Efficiency.* Proficiency in all subjects. Excellent in German; all-round ability.

Observation. Intense and not easily diverted.

Fatigability. Not easily fatigued.

Speed of Learning. Fast tempo.

Work Tempo. Works very fast. Is first to finish (except in arithmetic).

Memory. Memorizes very slowly but tenaciously. Surprisingly good in delayed recall. Particularly good in verbal memory.

Imagination. Oral and written composition comprehensively prosaic. Compositions unimaginative."

The vagaries of judgment, however, are still active

in reducing the margin of accuracy. The observational records lack nothing in concreteness, but the absence of numerical ratings contrasts unfavorably with the mathematical exactitude of the I.Q. and makes comparison very difficult. A scale that does not employ arithmetical or algebraic summation hinders appraisal in proportion to the number of neutralizing compensating characteristics revealed by it. The enumeration of qualities is a process of characterization requiring a final synthesis for its completion.

Dr. June Downey has developed a series of tests comprising a " scale for the measurement of the volitional pattern." The scale is composed of twelve subtests. It is an individual scale. The medium is handwriting. The subject is invited to write a sentence disguising and modifying his writing under conditions of constraint provocative of significant volitional and emotional reactions. Each test is rated on the basis of 10. They measure such qualities as assurance, flexibility, speed of movement, motor impulsion, resistance, tenacity, coördination of impulses, freedom from inertia, motor inhibitions, care for detail, speed of decision, etc. When plotted as a graph, they represent the " will profile " of the individual.

The tests are grouped in series of four, to correspond to three types of will profiles. A high score in speed of movement, freedom from inertia, inflexibility, and speed of decision points to the " quick, adaptable, fluidic type of person." Those who excel in motor impulsion,

assurance, resistance, and motor inhibition belong to the aggressive, willful type, while excellence in care of detail, coördination of impulses, perseverance, and revision indicates a " slow, accurate, tenacious type."

Dr. Downey's investigations led her to conclude that individuals with high intelligence rating but low volitional score do not usually fulfill the promise of the high I.Q., that low ratings in either test are rarely accompanied by high ratings in scholarship, but that inferior intelligence may be amply compensated for by the possession of superior volitional traits.

Dr. Downey has recently devised a group test called the Downey Group Will-Temperament Test.* It consists of thirteen tests, of which only twelve are used, in whole or in part, in estimating will-temperament. Some of these are highly ingenious. Take the tests for measuring noncompliance and volitional perseveration. Noncompliance, or the capacity to resist suggestion, is tested by means of a memory test and a test of high suggestibility. The examinee is confronted with sixteen statements relating to the conduct of the examination, which consists of the reading of a list of ten words containing many repetitions. Part of the test seeks to discover whether the examinee was critically observant of the examiner's actions, such as the manner of holding the card on which the words appear, the use he makes of the pencil clasped in his right hand, and the final dis-

* Copyright, 1920, by World Book Company, publishers, Yonkers-on-Hudson, New York.

position of the card. The fidelity of memory is attested by the statements marked " true " or " false." This is the preliminary step. Now comes the actual appraisal of the degree of noncompliance. Upon the completion of the test, the children are informed that " of the sentences you have just marked, *eight* are false, *eight* are true. Draw a circle around any ' true ' or ' false ' that you made a mistake in marking, and underline the other word instead. For example, sentence 2 reads, ' Two of the words in the list rhymed.' If this was false, but you underlined the word ' true,' draw a circle about the word ' true ' and underline the word ' false ' instead."

" Volitional Perseveration " is a test of perseverance. It is measured by the time consumed by the examinee in the attempt to write " United States of America " in a thoroughly practiced disguised hand, which practice terminates when in the judgment of the examinee the disguised handwriting has been perfected. Judgment and decision play an important rôle.

The tests here described typify the outstanding characteristics of the Downey Tests: (1) Success hinges upon the degree of whole-hearted coöperation the examiner can enlist. Unless the examinee cares to exert himself, the test itself is not sufficiently engrossing to evoke the incentives inherent in " question and answer " tests. (2) The examinee is called on to make highly unusual adaptations whose consummation may not indicate high adaptability in school situations. It is a

question whether the novelty of the tasks does not furnish the preponderant motive force.

The elaborate scoring device is an unmitigated evil. Scores are represented in the form of graphs and the assignment of values. Dr. Downey, however, holds out the hope of future revision in the direction of greater simplicity and objectivity.

Emotional and volitional traits are often vitiated by a low moral tone. Dr. Kohs has devised an Ethical Discrimination Test to throw light on the moral side of personality. It consists of six tests: (1) "Social Relations," (2) "Moral Judgment," (3) "Proverbs," (4) "Definitions of Moral Terms," (5) "Offense Evaluations," (6) "Moral Problems." The method is one of alternative statements only one of which is the correct one. "Social Relations" is subtitled "a test of common sense." Each of the twenty statements suggests social situations and three possible reactions. Question 5 states that:

> "Criminals are locked up
> to get even with them.
> to protect society.
> to make them work."

The examinee indicates the right reason by marking it with a cross. The test for "moral judgment" is borrowed from Pressey. From related "vices" arranged in groups of five, the examinee is asked to underline the worst. The "Proverbs" test is a modified

Otis test. "Definition of Moral Terms" is self-explanatory; the "definition," however, is limited to picking out the right word from a group of words, some of which are highly irrelevant. The "Offense Evaluations" test invites the examinee to indicate in tabular form whether the twenty-five offenses mentioned are deserving of praise, censure, prison, jail, death, or nothing. The "Moral Problems" test consists of a list of ten hypothetical situations of a high moral voltage, with the suggested solution for each. Problem 9 commands that:

"You should not take apples from another man's garden, because
too many apples give you a stomach-ache.
the apples are not yours.
the apples may not be ripe."

The author has published tentative norms and tentative classifications. The latter divide people into these classes: average adult, subaverage adult, mildly inadequate, inadequate, strongly inadequate, morally deficient, high-grade, middle-grade, low-grade, and moral imbecile. No attempt is made to evaluate ethical discrimination above normality.

The weakness of the Test inheres in its complete divorce between action and intention. Moral judgment operates to place morality on a higher plane, and moral problems are, in a way, dramatized moral acts, but no formula exists for transmuting ethical discrimi-

nation into ethical conduct. The biggest rogues could make a perfect score in ethical discrimination, for the will-to-do can be tested only by day-by-day conduct.

(1) *Limitations of Teachers' Estimates*

Teachers' estimates suffer from the following limitations:

1. Teachers are prone to underestimate the ability of a young child and to overestimate the ability of an older child. Age has a way of adding fresh accretions to its store of practical wisdom, so that teachers find it easier to multiply contacts with the older pupils.

2. Wittingly or unwittingly, teachers' estimates are colored by sympathy or antipathy. The troublesome pupil is penalized. A mark of C in conduct is apt to outweigh an A in mental alertness, and an A in conduct may palliate an unsatisfactory rating in proficiency.

3. In departmentalized schools, teachers frequently base their estimate on school marks alone. These are too often the product of chance, and fortune may smile on the laggards and frown on the diligent.

4. Teachers usually stress performance as against power. What a pupil does as evidenced by tests is given greater prominence than what he can do or does only sporadically. Therefore, teacher estimates throw no light on extra-intellectual determinatives of school success such as amount of effort expended to achieve a certain result, the rate of learning, the quality of learning, the aptitude for coöperative study, the relative

strength of the various mental "functions," and the nature of the pupil. Not that teachers disparage the influence of non-intelligence traits. They either intrust this information to an overburdened memory or find the curricular demands too exigent for exploration of individual minds. It is only the irrepressibly active pupils who impress themselves indelibly on the teacher's mind.

Factors that Enhance the Reliability of Teachers' Estimates. Teachers' estimates might be rendered more reliable if teachers would judge a pupil's ability in terms of (1) scholastic proficiency, (2) degree of applicability, (3) quality of memory, and (4) emotional stability. These are the preëfficients of school success; their exhaustive examination lends dependability to diagnosis and validity to prognosis; they determine the pupil's scholastic status quo and, in conjunction with the intelligence rating, indicate the probable direction of future progress.

(*a*) *Measuring the Scholastic Proficiency.* Scholastic proficiency may be measured by means of school tests or by means of standardized educational tests. The latter may be summarized in terms of educational age or accomplishment age. Both measure achievement, school tests stressing conformity with curricular requirements and educational tests emphasizing the acquisition of fundamentals. The former disclose the effectiveness of both the teaching and learning process; the latter appraise the pupil's mastery of salient processes

and principles. The school tests measure knowledge first and power only secondarily; the educational tests measure power first and only secondarily knowledge. Since knowledge and power are but two manifestations of the same mental capacity, teachers should employ school tests to check performance and educational tests to evaluate power.

(*b*) *Measuring the Degree of Applicability.* Diligence is not an intellectual trait, but it is a trait of personality that conditions mental activity. In a way, the degree of scholastic proficiency measures diligence, but the teacher must disentangle it from other contributory factors. This entails prolonged observation of every pupil throughout the school day. The teacher should note how quickly difficulties are surmounted, how doggedly students cleave to a line of action, how easily failure paralyzes their efforts, and how faithfully they obey the injunction " to try, try again."

(*c*) *Ascertaining the Quality of Memory.* Like diligence, memory may be measured indirectly. But careful classification of pupils demands a knowledge of the quality of retentive capacity — the fidelity of immediate memory, the tenacity in delayed recall, or what Burt calls " long-distance memory," and above all its relative efficaciousness in retaining numbers, words, and logical connections. It is also important to know whether memorization is rendered more durable by the stimulation of the sense of hearing or of sight; whether the pupil is a visile or an audile or perhaps a motile.

(*d*) *Appraising the Emotional Stability.* The emotionally stable constitute the class of pupils concerning whom prognostication is least disappointing. They are not subject to the debilitating effects of the recalcitrance of will or ebullience of temper; their intellectual efforts are not easily deflected by rapid alternations of moodiness and exaltation. The unstable child does not fit into a class organization, and the administrator must be placed in possession of the data of instability in order to institute adequate remedial measures.

(*e*) *The Superiority of the Composite Estimate.* The teacher's estimate based on character analysis is easily superior to the impressionistic judgment, which does not weigh psychological forces that reënforce intelligence, for intelligence is virtually the child's entire personality asserting itself in the conquest of environmental forces, not merely an aspect of mind activity, and such orientation of personality exalts teachers' estimates to the position of a criterion for judging the validation of intelligence tests.

An example should elucidate this point further. An 8B class was given the Otis Intelligence Tests. The class teacher was directed to arrange the class in order of standing, taking into consideration all school marks obtained both in class recitations and in class tests. There was a wide divergence between the intelligence scores and the teacher's estimates; the degree of correlation was very small, barely 40. It was

then suggested that the teacher disregard class marks and that he arrange the class once more in a seriation based upon pupil potentiality, *i.e.*, that each pupil should be placed in that rank-order to which his capacity for mental alertness and diligence, if exercised consistently and continually, would have entitled him. In fine, not performance, but promise was to be the touchstone. The correlation between the Otis scores and teacher's estimate now reached the significant figure of 82.

(2) *Shortcomings of Tests of Non-intellectual Traits*

Test makers have encountered serious difficulties in devising measures of non-intellectual traits. Temperament, morality, and motility do not lend themselves as readily to testing as do intelligence and its correlate, educational achievement. For one thing, their attributes are difficult to resolve into component elements. An act of intellection, especially in its scholastic aspects, stands out definitely and saliently. If other incalculable factors are involved, allowance may be made for them in the final evaluation. Predominantly emotional or volitional behavior has its roots deep in the soil of instinct and habit, and it becomes difficult to trace its ramifications and labyrinthine windings in the unconscious. Again, the learning process and its products do not fluctuate under changed circumstances; the mind learns and stores this learning in fairly stable ways. On the other hand,

a sudden change in the visceral rhythm, an alteration in social outlook, deterioration of bodily vigor places a new complexion upon the individual's emotional and moral reactions. We saw that it is very hard to fixate intelligence in its permanent aspects through the instrumentality of an intelligence test; it is almost impossible to do so in other dimensions of personality.

Most tests of non-intelligence traits have not surmounted this difficulty. The subjects have been confronted with laboratory or selected situations — situations from which the elements of distraction and personal gain are absent. The subject's responses are modified by (1) desire to do what he thinks he ought to do, (2) desire to do what he thinks the examiner wishes him to do, (3) desire to make an objective, impersonal, and magnanimous decision. The situation-response mechanism becomes intellectualized and idealized. The subject appears at his best. What the test fails to evoke is the natural situation and the natural response — conditions of spontaneous behavior.

May and Hartshorne studied this desideratum by analyzing the varieties of situation-response combinations operative under testing conditions. The most important, for testing, are the following:

	Situation	Response
1.	Natural, uncontrolled	Natural, undirected
2.	Natural, controlled	Natural, undirected
3.	Natural, controlled	Natural, directed
4.	Experimentally controlled	Experimentally directed

Type 1 represents conduct uninfluenced by either the testing material or the examiner. The nearest approach to it is the handwriting and composition scales. The pupil works under natural conditions; the testing is done by the teacher upon the product and not upon the pupil; and there are no subjective concomitants of direct testing. In the sphere of non-intellectual traits, the report card or the observational record possesses the same characteristics.

Type 2 represents conduct which the individual exhibits when acting under guidance. The period of childhood furnishes many examples. The child is always under some form of control, in the form of precepts, admonitions, and direction. The response may take place far from the prying eyes of parent or teacher. The response is undirected and is, in a way, a measure of the efficacy of parental and tutorial guidance. Test makers have attempted to approximate such behavior by controlling the stimulus and learning the response unfettered. Hartshorne, for example, arranges books, pictures, and articles in a room in a certain manner, puts the subject in it, and observes — unnoticed — his behavior; in this instance, his idea of how a room ought to look. Voelker, wishing to test the subject's honesty, sends him on an errand with 25¢, arranging with the storekeeper to give more change than the transaction calls for. Or consider this test, whose source is unknown. A group of boys went on a hike. A game was played on the way so planned that, when finished,

each boy had a paper in his hand bearing his name. The motive of the game was to find out which of the boys would dispose of the paper in refuse cans conveniently placed near by. In all these bits of behavior, conduct approximated natural settings, *i.e.*, the subject acted as it was his wont to act. The situations, though controlled, were natural — they were fragments of life.

Type 3 is exemplified in the school examination. The situation is natural; the pupil has become habituated to this form of recitation. The response is natural; the pupil sets about answering the questions, employing the technique mode habitual by long practice. Voelker and Allport have devised such tests.

Unfortunately, very few standardized tests conform to the above-mentioned types. Rather do they follow Type 4. The subject is asked to do something that he never encounters ordinarily, in a manner that is just as alien to everyday existence. As May and Hartshorne put it: "The chances are not one in a lifetime that a child would be confronted with a piece of paper containing lists of five words each and told to cross out the worst word in each list, except under test conditions. This thing probably never happened in just this way in the history of the world until Pressey invented his test."

If the testing technique is inadequate to evaluating non-intellectual traits, what means may be employed? It would seem that our main dependence must be placed on (1) observational records, (2) trained judgment of

those who come in intimate and prolonged contact with the subject, (3) scales of personality traits, and (4) psychographs. Allport sums up trenchantly the case against character tests: "Character is truly revealed only in the vital issues of life. It will probably be some time before we shall have advanced beyond the methods of the business man who, desiring to employ an office boy, allowed each applicant to walk into a room where lay an unguarded pocketbook while the prospective employer watched behind the crack of the door."

B. Studies in the Nature of Character

Hartshorne and May have recently published one of a series of studies on the nature of character. It is entitled *Studies in Deceit.** This investigation was made in coöperation with The Institute of Social and Religious Research, and is composed of two books: Book One describes general methods and results, and Book Two gives the statistical methods and results.

The investigators have spared neither trouble nor expense and have produced a conscientious statistical work which bids fair to become an inexhaustible source book for further research in the field of character. They have restricted their activities to the trait of honesty and its correlate dishonesty because these types of behavior lend themselves readily to quantitative evaluation and are of profound social significance.

* The Macmillan Company, 1928.

(1) *Types of Deceptive Behavior*

The authors constructed and adapted tests for measuring (1) the cheating type of deceptive behavior, (2) the stealing type, and (3) the lying type. Since some of the measuring material is not strictly tests, the word " techniques " is employed as a convenient descriptive term.

Methods for Measuring the Cheating Type of Deceptive Behavior. The authors have restricted themselves to measuring cheating as exhibited in classroom situations, in work done at home, in parlor games, and in athletic contests. Under classroom situations are to be found the following techniques:

(*a*) *The Duplicating and Multiple Testing Techniques.* Three kinds of testing material make use of duplication. All are of the short-answer type of test. The papers are collected and duplicates made. Their contents do not differ from those used in intelligence tests. Later the test papers are returned together with answer sheets or keys. The subjects score their own papers. Those who cheat do so by copying answers from the key, thus decreasing their scores. Tests (1), (2), (3), and (4) were supplied by Professor Thorndike and comprise the *Double-Testing Technique.*

(1) *The Information Test.* The subject is asked to reply to all of the twenty-eight questions by encircling the answers in ink. The pupil inclined to cheat must erase the circle and draw another one.

(2) *The Sentence Completion Test.* Fifty-five incomplete sentences are given and the pupil can cheat by adding more items or by changing words previously written in pencil through erasure or rewriting.

(3) *The Word Knowledge Test.* There are 120 items of the multiple-choice (5) type. The number of the preferred answer is placed at the margin on a dotted line. Erasing the number and substituting the correct one or writing down more numbers after the test papers are collected constitutes cheating.

(4) *The Arithmetic Test.* It embraces fifty examples whose answers are written at the margin. Cheating occurs in the same way as in test (3).

(5) *The Thorndike-McCall Reading Scale.* Opportunities for cheating are obvious.

(6) *An Original Disarranged Sentence Test.* There are fifteen word groups all of which the child is urged to answer. A line drawn through a blank space indicates an omission. Cheating is easily practiced.

(7) *An Original Spelling Test.* There are ninety words some of which are misspelled. Only misspelled words are to be checked. Cheating is readily detected.

(*b*) *The Double Technique.* Five kinds of tests were selected from Thorndike's intelligence level material: (1) arithmetic problems, (2) mutilated sentences or sentence completion tests, (3) information test elements, and (4) word knowledge or vocabulary test elements. These are called I E R Achievement Tests; they were developed by the Institue of Educational

Research. These tests have been validated as intelligence testing material. The difficulty of each element has also been experimentally determined. Two parallel forms of them exist. The I E R tests have been standardized for children from grades 5A to 9A. Thus the I E R Achievement Tests yield a measure of intelligence in terms of a point score and at the same time test deceptive behavior under classroom conditions, and achievement in certain subjects.

(*c*) *Triple-Testing Technique.* Technique (*c*) is made up of speed tests. The authors made use of some older psychological speed tests with a view to lessening cost, increasing reliability, and widening their application so as to include all grades. Six tests were finally selected.

(1) A simple test of addition requiring the rapid addition of one or two digit combinations, such as 4 and 5, 6 and 2.

(2) A number checking test similar to the one in Army Beta.

(3) Cancellation of A's.

(4) The Digit Symbol Substitution Test (found in the National Intelligence Tests).

(5) Making dots in small squares.

(6) Cancellation of single digits as in the Woodworth and Wills series.

The Speed tests are given three times. This is really the Triple-Testing Technique. The first two are given as achievement tests without extensive opportunities for cheating. They are given as practice tests.

Cheating is facilitated on the third trial when self-scoring is invited.

(*d*) *The Improbable Achievement Technique.* The assumption underlying this technique is that when the improvement resulting after the performance of self-scoring widely exceeds the normal ability, deception is unequivocally indicated. The amount of divergence and its statistical significance will be explained later. Two kinds of tests are used, tests using objects and tests using paper and pencils.

(1) *Puzzle Performance Tests.* Three puzzles are used: (*a*) The Puzzle Peg (or Peg Solitaire), (*b*) The Fifteen Puzzle, and (*c*) The Weight Discrimination Test. The first two are toys containing considerable puzzle element, and consisting of parts which, when properly manipulated, result in their sequential arrangement, culminating in a "solution." The child is provided with a score sheet on which he records his performances. There is no way of knowing whether deceiving subject falsified performance and made a true record, or whether he merely cheated by falsifying the record. In one case a checkup was made and it was found that "in over ninety per cent of the deceptive cases the peg boards and score sheets agreed."

The Weight Discrimination Test is a Binet test. On the first trial pupils were asked to turn the numbers down and arrange the pill boxes in the order of weight. On the second trial they were told that the correct arrangement was 1, 2, 3, 4, 5, 6, 7 and were invited

to repeat the operation without looking at the numbers.

(2) *Paper and Pencil Tests, or Coördination Tests.* This battery comprises three tests: (*a*) a Squares test (Cady), (*b*) a Circles test (similar to Voeker's and Cady's), and (*c*) a Mazes test (suggested by Cady's). The opportunities for cheating are provided in the invitation to " peep " while directing the pencil through the ramifications comprising (*a*) and (*c*).

(*e*) *Methods for Measuring the Cheating Type of Deception outside the Classroom.* The techniques and tests described measure the cheating type of behavior as exhibited in classroom situations. Other techniques have been evolved for measuring cheating in (1) work done at home, (2) athletic contests, and (3) in parlor games — all of which are situations natural in point of interest for the subjects tested, and in the matter of social environment. In (1) the pupils were given the I E R Word Knowledge Test, Form T, and were told to finish it at home. They were cautioned twice not to seek help. Form II was given later in class along with the other I E R tests with no answer sheet. In (2) four tests of physical ability were given: (*a*) the dynamometer test, (*b*) the spirometer test, (*c*) the chinning test, and (*d*) the standing broad jump test. The pupils were asked to attempt each test. The examiner handed each one a score card, demonstrated each instrument, invited the pupil to make one or more practice trials, and exhorted him to try his best. He then left

the room. The unsupervised trials recorded on the score cards were then compared with the practice trial (which the examiner had noted surreptitiously) for evidence of cheating.

In (3) a party atmosphere was first evoked. The games consisted of (*a*) a peeping stunt (a modification of the coördination material); (*b*) pinning the tail on the donkey, or the arrow on the target (Under the conditions in which the games were played no child was likely to achieve the task without "peeping."); and the (*c*) bean relay, a modified potato race. Each child was given an opportunity to make at least eight runs. His score was the descrepancy between runs and beans.

Another game was that known as the Mystery Man, which could be used both as a test of cheating and of stealing.

Methods for Measuring the Stealing Type of Deception. Space is not available for listing all the methods. We will content ourselves with one sample test: The Planted Dime Test. In connection with the puzzle tests each pupil was given a little box containing several puzzles not all of which were used. In each box was a dime " ostensibly belonging to another puzzle which the examiner showed to the pupils but did not ask them to solve. This other puzzle required the use of a dime, but no mention was made by the examiner of the dimes in the boxes." The boxes were numbered and distributed according to a seating plan. The pupils were

directed to return the box to a large receptacle and a check was later made on those who abstracted the dime.

Methods for Measuring the Lying Type of Deception. The type of lying measured by the investigators is described as that form of deceptive behavior which consists in " making a false statement instead of a true statement when the truth is known to the one making the statement." This situation is uncomplicated by intentions to deceive and intentions are not easily discoverable. To measure " lying " adequately the authors have analyzed this form of deceptive behavior into (1) lying to escape disapproval, and (2) lying to win approval. The former was measured on the basis of a Pupil Data Sheet containing a mixture of unconscious and incriminating questions, given to each pupil in connection with the I E R tests.

Lying to win approval was tested by an attitudes test of which two forms are available. They contain 36 questions of the " Yes " or " No " variety.

The Personal Data Sheet was checked against the known proclivities of the subjects and indexes of dishonesty were assigned to each case. The attitudes test was scored on the basis of answers furnished by adults who placed themselves imaginatively in the shoes of youngsters. They brought to their aid the evidence cast up by retrospection and introspection. The statistical manipulation of the score will be discussed in the next section.

(*a*) *Methods of Scoring.* In scoring the tests, the

investigators have used the *amount score* and the *fact score.* The amount score is the number of changes made when dishonesty is resorted to. Any single change is counted only when the change affects the score; two changes are counted. Pupils may change the spelling or wording without changing the answer. Fact scores or simply " c " (cheating) are earned whenever any cheating has, in all probability, taken place. In the Puzzle Tests, the amount score for Squares equals number of corners turned correctly. Fact score allows 13 corners; 14 or more is called a " c." The same procedure is followed in the Circles Test and in the Mazes Test.

In the formulation of norms for fact scores, the authors made use of statistical figures and computations to which reference is invited in the chapter on Achievement Tests. These will be discussed under Standardization, and Interpretation.

(1) *Standardization.* The unit of measure was selected with a view to serving as a common denominator for so varying a series of tests. The authors argued that in all measurements two things are required: a unit of measure, and a point of reference. In the realm of the physical sciences we have the inch, the yard, the mile; the point of reference such zero points as sea level for altitude and the freezing point of water for centigrade temperatures. In dealing with statistical data, the point of reference might well be the mean and the unit of measure the standard deviation (S.D.) or " Sigma."

Applying this point of reference and its correlative unit of measure to such data as "Typical Gains and Losses in Arithmetic between the First and Second Day When No Cheating Was on" they found the mean to be 1.06 and the standard deviation, 3.10. Now Case G, who lost 8 points, lost a total deviation 8 plus 1.06. or 9.06. This deviation in terms of S.D. becomes 2.92 S.D's (9.06 ÷ 3.1). The authors have named the common measure the XI score or amount score.

The scores on the Pupil Data Sheet were standardized in terms of an honesty index or a lie index. The technique for establishing the latter is very complicated and interested students may best turn to the book.

(2) *Interpretation.* Since the authors' concern was with deceptive behavior it became evident that a line must be set beyond which deviation from "honest" behavior would indicate unequivocally the practice of deceit. In order to establish the boundaries between "honest" deviations due to chance and dishonest behavior, the authors fell back on the properties of the normal curve. The base line of this curve when divided into S.D. distances (the authors restrict themselves to 6 S.D's. ranging from − 3 S.D. to + 3 S.D.) encloses most cases (99.73%). In accordance with this conformation, there are 15.866 in 100 chances that a deviate of greater than − 1 S.D. will occur. In the case referred to above it was found that the child's deviation of 8 points in the Arithmetic test of the I E R material was equivalent to 2.92 S.D's. Now, the table of probabilities

shows that but two cases in a thousand will exceed this deviation. Similarly, a deviation of 9 points on the same test is 3.24 times the standard deviation of the honest Arithmetic differences and is less likely to occur. It is, therefore, well within the realm of probability that the pupil whose score registers a difference of nine between a performance under " honesty " conditions and one under " dishonesty " conditions, has behaved dishonestly. Specifically, the aspersion of cheating will be unmerited in six out of every 10,000 cases.

In like manner, but not to the same degree of statistical verification, the authors have established cheating norms for all techniques. In the I E R tests the limits are 10 for Completions, 9 for Information, and 22 for Word Knowledge. In the Speed Tests, it is 13 for Additions, 12 for Number checking, 14 for Underlining A's, 24 for Digit symbols, 45 for Dots, and 18 for Underlining Digits. In all instances the " dishonest " difference is three times the standard deviation.

(2) *Practical Value of the Deception Tests*

The Tests as Instruments of Social Study. Having studied quantitatively a range of deceptive possibility evocable by the varying techniques perfected, the authors set about to discover what relations deceptive behavior sustains to causative and provocative factors within what may be called the total situation. Since deceit and its opposite is part of the concept of character, they explored the boundaries of honest and dishonest

behavior as they manifest themselves in specific situations arising from the interaction of the child's intimate environment and all discoverable antecedent potent and prepotent elements. This consideration would involve the study of antecedent factors, conditioning factors, and the specific situation. Phrasing it differently, the need was felt for a valuation — quantitative as far as possible — of heredity, attitudes, and opportunities. We may indicate their interrelation diagramatically:

Antecedent Factors	Intelligence Emotional Force Emotional Stability Suggestibility	Heredity
Conditioning Factors	Social Conflicts Socio-economic Background Cultural Background Moral Education	Attitudes
Specific Situation	Opportunities for Deceit Facilitating Influences	Behavior

For the study of these relationships, associations, and concomitants they invoked statistical devices. One was the composition of the distributions yielded by the various techniques. Thus the authors summarize the results of one technique by stating that they had tested 143 eleven-year-old boys with the I E R Intelligence Tests and found that the original scores range from 53 to 176, that the median or fifty percentile score is 102.4, that the 25 percentile is 86.7, the 75 percentile is 118.2, that the average or mean score is

103.3, that the semi-interquartile range is 15.75, and that the S.D. is 23–24. Against such a statistical description of the group it is a simple matter to define the status of an individual. To illustrate the method the case of John Doe who is twelve years old and is a member of the seventh grade is cited. To compare John's achievement on the intelligence test with that on the arithmetic test, it is necessary to know the mean and S.D. of the twelve-year-old seventh grade scores in these materials. Suppose John received 160 on the intelligence test and 60 on the arithmetic test and the group mean in intelligence is 150 and on S.D., 25, while the group mean and S.D. in arithmetic is 50 and 5 respectively. John's intelligence status is $\frac{160 - 150}{25} = \frac{10}{25} = +.4$. In arithmetic, $\frac{60 - 50}{5} = \frac{10}{5} = +2$. Therefore, John has acquitted himself much more creditably in the arithmetic test.

Another method of group and individual relationship makes use of correlation, and their correlates, partial correlation and intercorrelations. The former need not detain us here. The others will bear extension because the authors support their main conclusions on them. In order to find out whether the presence of a positive correlation between moral knowledge and conduct argues for a causal relationship or whether a third concomitant, say intelligence, links and binds them together, recourse is had to one of several formulae

— too complicated for any but statistical experts — on which the average intercorrelations between conduct, knowledge, and intelligence for different groups is substituted. But this procedure involves three variables. Therefore, one of them, in this case intelligence, must be kept constant. That is, only the scores of these pupils are considered who attain approximately the same intelligence score. In like manner one may have two variables, age and intelligence, but these must be kept constant. Technically the process of keeping variables constant is known as " partialing out."

Age, sex, intelligence, and health scores were obtained from the aforementioned techniques. Cultural background was measured by the Sims Score Card which yields a general socio-economic index.

The Burdick Apperception Test was used to explore the cultural background. The test consists of several hundred items which indicate the type of literary contacts, concerns the child's social aptitudes, demands general information, reveals personal relations within the family, states hypothetical situations for which the pupil must furnish a solution, and gets at the quality of the home furnishings. The method of scoring is rough; it consists of a comparison of answers of populations which differed widely in social background.

The authors subjected the data found by the use of the techniques and supplemental tests to thorough statistical scrutiny. They obtained data on the relationship between (1) sex and age, (2) intelligence and

cheating, (3) physical condition and cheating, (4) emotional condition and cheating, (5) the Sims Score and deceptiveness, (6) cultural background and deception, (7) race, nationality, religion, and cheating, (8) school status and school honor, (9) school deportment and cheating, (10) school methods and school morale, (11) systems of moral education and deceptiveness, (12) motivation and deceit, (13) miscellaneous concomitants and deceit, and (14) ideals of conduct and deceit.

Space is lacking for the examination of the arguments and supporting facts for each of the fourteen derivative studies. Part of item fourteen only will be elaborated upon because it exemplifies the authors' procedure at its best and because it yields arresting conclusions; the other studies but confirm the *a priori* and experimental findings of other investigators.

(3) *The Doctrine of the Specificness of Conduct, Attitudes, and Motivation*

The Specificity of Conduct. This doctrine affirms that the driving force in ideals of honor and honesty is insufficient to function automatically in all situations in which deceitfulness may be alluring; that the individual's honesty or dishonesty depends primarily on the specific situation with no carry over unless (1) situations have common elements, (2) honesty or dishonesty with their accompanying implication or consequences have been specifically associated with each situation. Three sets of supporting facts are presented: correla-

tion between tests; comparison of tests with respect to amounts of cheating; the distribution of cheating scores.

(*a*) *Correlation between Tests.* 1. Repetition of identical tests under indentical conditions after a lapse of one week yielded the same average correlations between the first cheating score and the second: for speed tests, .569 and for coördination tests, .566. 2. Two tests under the same technique but employing different material gave an average correlation of .440 for speed and .462 for coördination. 3. Correlations running from plus .400 to minus .003 were obtained when one test representing one type of situation was compared with another representing a different situation. Deviations in correlation resulted from differences in the situations. 4. According to Table CLXVIII * — "the *r*'s (correlations) perhaps signify in each case the extent to which the situations have elements in common. There is much in common between any two I E R tests or any two speed tests (*r* equals .527). The only differences are differences in material. But the speed and I E R tests differ from each other not only in the nature of the material but also in the opportunity to deceive. There is less in common between them. There is still less in common between the classroom situation and the home situation. Thus the correlation between two deception tests depends on the common elements between the two situations."

(*b*) *Comparison of Tests with Respect to Amount of*

* Book Two, page 212.

Cheating. 1. When tests differing in test situations are repeated they provoke different amounts of cheating. 2. Tests of the same situation type using different material do not have as wide differences — though they are statistically significant — as tests representing different situations. For example, in the coördination tests (squares test, circles test, and mazes test), we have almost identical situations. There is more cheating in the mazes because " deception is a matter of the nature of the situation as much as, if not more than, the nature of the child."

(*c*) *The Distribution of Cheating Scores.* Unified traits manifest themselves in behavior which when evaluated quantitatively approach the U-shaped curve of distribution: it is bimodal. If honesty is a unified trait, a child would be uniformly honest or dishonest. Yet from Table CLXII,* under the column headed " Per Cent," it is evident that 3.2% of a population of 2443 uniformly cheated and 7% did not cheat once. Here there is unmistakable evidence of the presence of a unified trait. But of the other 90%, some cheated once and others nine times in every ten opportunities. This condition does not favor the existence of a unified trait.

(4) *Evaluation of the Authors' Techniques*

It is difficult to employ temperate language in praising *Studies in Deceit.* It is equally difficult to appraise the influence of the techniques developed upon future

* Book Two, page 220.

research. We have maintained consistently that statistics is a highly explosive force which only experts can handle with safety. Expertness, however, is not merely a matter of mathematical skill. Rather is it a matter of scientific attitude. For the danger lies more in interpretation than in manipulation. We have had a horrible example in the injudicious use of the data yielded by the Army Tests. Laymen and educators accepted the validity of both tests and techniques and proceeded to twist and distort the findings in every conceivable way. The word "moron" gained tremendous vogue and was bandied about indiscriminately as a classificatory term and as a term of opprobrium. Racial snobbery grew rampant throughout the land and social antipathies were violently stirred.

The same fate threatens to overtake the findings of Hartshorne and May. Their investigation of the interrelationship between proneness to deceptiveness and various prominent religious denominations may arouse bitter animosities although the authors qualify their conclusions in ways calculated to soothe the susceptibilities of any right-thinking individual. Yet, to many minds, there is a provocative suggestion of emphasis in the most carefully worded understatement. Perhaps it would have been better if the authors had postponed the account of their research into the moral influences of religion as well as other " partial " studies against the accumulating evidence which they are preparing for future publication. There is enough intellec-

tual dishonesty abroad to warrant the entertainment of the gravest fears that many of their tentative conclusions, stated with a wealth of restrictive clauses, will serve as foundations for highly precarious structures, educational and social.

But from a strictly psychological point of view, *Studies in Deceit* may well be styled epoch-making. The authors are to be commended for their meticulous care in formulating criteria for the testing material and for their finesse in the statistical elaboration of the scores.

As instruments for measuring manifestations of character, the tests and techniques satisfy all scientific requirements of a good test and surpass in validity most tests in the field of intelligence.

As to the doctrine of the specificity of conduct: We believe that it may well prove less irreconcilable than it appears. Perhaps agreement between the doctrines of specificity and unified traits may be effected by analyzing the latter. The child progresses from the sensori-motor level to the next higher level, from habits to ideals. If we wish to break the "I ain't" habit we begin with specific drills on the right response but reënforce it with the ideal of cultured diction. Each supports the other. Moral situations are more complicated. We are balked in our attempts to abstract honesty from all its manifestations by the antagonisms created by the prevailing looseness of accepted codes. It is not the unvarying reaction to

reprehend such conduct as cheating the conductor or the gas company or the government. Honesty, then, becomes for many a highly flexible concept; and the lack of universality undermines its potency.

Again there exists a great deal of moral ignorance. Children do not know that many of their habitual reactions are unmoral: they see others practice behavior which is not strictly honorable without arousing condemnation. Besides, every situation eliciting moral behavior is so closely hemmed in by personal considerations, pertinent attitudes, and the fiercely impelling urge of emotional sets that it rarely obtains high intellectual visibility: the controlling, comparing, and generalizing intellect is rudely shouldered aside. The opportunities for generalizing conduct are less abundant than those for generalizing knowledge. Ideals of conduct, such as the ideal of honesty, mature very slowly and unlike purely intellectual concepts, must arise from a prolonged practice in the acts that constitute them.

There is, therefore, little carry-over, first, because every such concept is built out of a multitude of discrete acts; secondly, because the dominant force, the emotions, are eternally opposed to the exercise of intelligence. There results, then, the enthronement of the situation as the arbiter of moral behavior, and its characteristic of inducing specificity of conduct. It is our conviction that a deeper understanding of the true inwardness of ideals may serve to abate the antag-

onism between the doctrines " specificity " and " unified trait," for the force of emotionality loosed by a moral situation renders it, to a large extent, decisively specific.

C. Summary and Conclusion

(1) *Present Status of Testing Movement*

The testing movement is slowly delineating itself as an attempt to measure, by standardized tests, the intellectual capacity of human beings. By far the greatest strides have been made in measuring the intelligence of children. These tests are based on the monumental work of Binet and his collaborators, who have evolved types of test problems drawn empirically from the social and educational environment of civilized society and adapted them to the mental capacity of children. The responses culled from experimental groups of a wide range of mentality have been supplemented, statistized, and scaled by other investigators. The results have been summarized in the I.Q., which assigns to tested persons their level of intelligence in terms of mental age and chronological age.

1. Psychological Criteria. Devisers of tests are guided by the following psychological criteria :

1. Tests should contain a high degree of validity. A test is not valid unless it conforms with the findings of other ascertainable measuring instruments, such as accepted tests or trained teacher judgment.

2. The test should be as independent of schoolings as possible. The child who enjoys superior educational

advantages would naturally score higher in those elements of the test which are based on school knowledge. On the other hand, Fox voices the warning that, " since ability and maturity *in vacuo* are empty conceptions, the attempt to measure native intelligence or ' mother wit ' apart from the material on which it is exercised is seductive but impracticable."

3. The examinee must be assured a degree of physical and mental comfort attainable only in an environment free from physical discomforts and the stress of the examination-complex. To this end, the examiner should surround the child with all facilities of a well-equipped classroom and inspire him with a feeling of attentive ease.

4. The range of difficulty should be such that, " if 50 per cent of the group tested, or even 20 per cent, make zero scores, the test is unsatisfactory as a measure of a wide range of intelligence."

5. Intelligence being but one aspect of the personality-complex, the prognosticative value of an intelligence test is in proportion to the appraisal, by subsidiary means, of other factors of personality, such as emotional stability and temperamental bias.

2. Criteria for Choosing Test Material. The choice of material has not been subjected to experimental investigation so that we are without dependable data as to the psychologic significance of the questions comprising the tests of intelligence. However, devisers of tests are not without workable criteria which are being

formulated with varying degrees of scientific accuracy. The findings of A. S. Hamid, published in the *British Journal of Psychology* for October, 1925, indicates the trend of such experimentation. The author chose 793 individual test questions, with which he tested 373 subjects, adults and children. These were selected from a much larger group on the basis of records and teachers' judgments and embraced the brightest and dullest with a sprinkling of those of medium ability. Tests which correlated highly with the rankings established by the records were subjected to closer study with a view to appraising their merits. Hamid found that:

1. Tests of information are of doubtful diagnostic value, only 11 being classed as good and 45 as poor.

2. Catches are good provided the logical relationship is not obscured by obliqueness of expression. The following can be readily solved with careful reasoning:

   Train is to station as ship is to . . . . port, sea, shelter, destination.

3. Very easy tests have no value:
   Ounce is to pound as inch is to . . . .

4. A moderate degree of abstractness is good.
   Low degree of abstractness: Man is to skin as tree is to . . . . branch, leaves, root, bark.
   High degree of abstractness: Money is to finance as labor is to . . . . wages, workmen, industry, factory.

5. Too finely drawn distinctions are bad:
   Loss is to sorrow as gain is to . . . . profit, advantage, mirth, enjoyment.
6. Inept analogies are bad:
   Couch is to chair as luxury is to . . . . comfort, wealth, stool, pleasure.
7. Loose analogies are unacceptable; they create the possibility of several choices:
   Duet is to orchestra-playing as partners are to . . . . whist, company, business, finance.

Since nonscholastic content is a desideratum, preference should be given to those tests which diverge as widely as possible from the traditional subject matter found in school curricula.

3. Administrative Criteria. The administrator is actuated by motives other than those that animate the psychologist. He wants to know what tests yield diagnostic and prognostic data in the most economical, dependable, and expeditious manner. From this point of view the following items require consideration:

1. The cost of the material should be moderate. The present tendency is to reduce costs, especially on large quantities.

2. The simpler the testing procedure, the better. Administrators are reluctant to enter upon a testing program requiring highly trained examiners.

3. The complexities of departmental changes and comparatively short periods of instruction tend to bring into favor the shorter tests.

4. The more compact the test, the greater its popularity.

5. Considering the extreme diversity encountered in the matter of marking, recording, and interpreting findings of tests and the variety of supplementary explanatory pamphlets and record blanks employed, administrators prefer tests which adhere to standardized categories of classification and which combine into one booklet all information and forms often consigned to bulky manuals, graphs, and numerous sheets.

6. The fewer instructional pauses, the greater the opportunities for continuous concentration and the less provocation for breaches of discipline. Otis' Self-Administering Tests are admirably adapted to subserve both requirements.

7. A predominately linguistic test should be followed by one of material appeal. The National Intelligence Tests should be checked by Dearborn's Test.

8. The closer the kinship of the experimental group responsible for the norms with the school population to be tested, the greater the probability of success. The National Intelligence Tests give tables of varying norms derived from socially homogeneous samplings in many sections of the country.

9. Whenever the combined teachers' judgment is at variance with the findings of the test, the pupil should be tested through individual tests of both educational intelligence tests and personality tests. Where harmony cannot be established, teachers' judgment should prevail.

10. The I.Q. derived from group tests should not be accorded the same consideration bestowed upon the Stanford I.Q.

11. The greater the number of forms of the same test, the more useful.

12. The less elaborate the answer forms, the less fatiguing for the teachers charged with marking papers.

13. The administrator who is beguiled by the apparent simplicity of the " instructions " into omitting the minutest elucidation in conference is jeopardizing the entire testing program. Teachers are flagrantly remiss in following directions and in adhering closely to the printed instructions.

14. A checking system should be installed to guard against (*a*) carelessness in marking, (*b*) errors in addition of subscores, (*c*) discrepancies between the I.Q. and the chronological and the mental age, (*d*) confusion between the I.Q. and the total score when the latter does not exceed the prevailing range of quotients, and (*e*) interchange of the chronological and mental ages.

15. The fact should be impressed on teachers that the I.Q. is but one measure of the learning personality, whose scholastic success depends on factors at present but obscurely discerned by the tester.

(2) *Directions of Needed Improvement*

The measurement of intelligence and non-intelligence traits are in need of improvement in the following directions:

1. Emphasis should be distributed equally between tests of intelligence traits and tests of non-intelligence traits. Teachers and administrators have been beguiled by the glibness with which some of the makers of intelligence tests recount the infallible virtues of their wares. While the ranks of devisers and popularizers of tests have not been invaded by quacks and charlatans, there are far too many instances of pretentious claims and facile promises. True, every movement is bedeviled by a fringe of amateurs, who, in the first blush of enthusiasm, accept conclusions uncritically and influence lay members to emulate their example. Some even surpass them in zeal, with the result that we are constantly harassed by trusting teachers who have been painfully disillusioned by intelligence tests.

The complete rehabilitation of intelligence testing can be attained only by disseminating the information that intelligence, *per se*, is a mere abstraction; that the teacher is not engaged solely with a thinking machine, but with a complex personality that thinks and feels and wills while undergoing instruction; and that testing intelligence is but measuring one aspect of personality, leaving out of calculations non-intelligence traits which contribute as weightily to the learning process.

2. We must strive for the standardization of the I.Q. Someone has remarked trenchantly that " figures don't lie but those who make them might." This indict-

ment might be made, with lesser force, to the practice of accepting I.Q.'s as of the same significance and omitting to inquire concerning the nature of the tests in use. An I.Q. derived through the use of the Stanford Revision is one thing, and an I.Q. based on a group test is quite another thing. The situation is still more complicated whenever I.Q.'s derived from various tests are averaged. A step in the right direction has been taken by the authors of the National Intelligence Tests, who qualify the I.Q. as an N.A. I.Q.

3. The need for simplifying the statistization of scores and norms has been emphasized in a previous chapter. Too many psychologists mistake statistical treatment of figures for predictive evaluation of answers to tests. Statistics, at best, is the handmaiden of psychology and not its master, and there are many imponderable elements in personality that baffle the most astute mathematician. Besides, traits of personality are too dynamic to be treated statistically.

4. The time has come to effect larger syntheses of tests. Laborers in the field of intelligence testing have worked independently of those investigators in the field of non-intelligence tests. We are faced with the anomaly of tests of intelligence that are in part tests of non-intelligence traits and tests of non-intelligence that are intellectualistic in their content.

Moreover, even those who recognize the usefulness of a complete appraisal of the pupil's personality have done little to combine the two into a testing program

calculated to study all its facets. It would constitute a step in advance to have devised an all-round test, within the compass of one booklet, containing approved forms of both intelligence and non-intelligence tests, retaining the designation I.Q. for the former and psychographic or other summation for the latter.

5. Coincidentally with the construction of integrated tests of intelligence and non-intelligence traits, a movement should be set on foot to impress upon school officials and the public the advantages accruing from exhaustive psycho-physical examinations of children. A complacent and tax-ridden public should be aroused to the hit-or-miss policy pursued by our schools in their war upon repeaters, delinquents, and incorrigibles. If tests possess any virtue at all, their use as instruments of adjustment and classification should be placed on a parity with other recognized pedagogic technique and not left to the whims of the principal or teacher or the frenzied zeal of the amateur psychologist.

An exhaustive testing program should include:

A. The assignment to each school of a trained psychologist versed in the technique of testing, both as an examiner and as an interpreter of results

B. A medical examination of each pupil upon his admission to school

C. Testing each pupil to determine:

(*a*) Level of intelligence

(*b*) Trends and dispositions revealed by tests of non-intelligence traits

(*c*) Scholastic ability as soon as his school status warrants it

D. Retesting each pupil at intervals of a year or two to note:

(*a*) Degree of improvement or retrogression

(*b*) Emergence of latent talents, skills, or interests (to reclassify pupils who " bloom late ")

(*c*) Physiological and anatomical changes such as height, weight, and expansion of chest, the onset of puberty and the rate of ossification. This information would act as a corrective of the other measures of personality, especially useful in instances where retardation may be attributable to physiological disturbances or physical retardation.

E. Such flexibility of school organization as would permit the ready transfer of pupils who display a sudden burst of energy or unequivocal symptoms of retardation

6. The test material embodied in the tests should be subjected to experimentation to determine its suitability. We cannot uphold our claim to scientific accuracy as long as the whole structure of testing rests on the foundation of material chosen empirically.

Moreover, the test material should include graded problems and devices calculated to test the capacity for configurative learning, such as learning by accretion, by assimilation, by gradation, and by redefinition. We might build up batteries of tests based on each type

of learning. For example, complete tests of ingenuity might be constructed to discover how well the examinee can put together what he already knows (assimilative learning), how successful he is in creating new wholes. Tests of learning by gradation may be devised in which a pattern must be completed. Thus:

15	25	30	40	45
24	18	13	6	—
18	6	36	—	48

The capacity for redefinition may be measured by equivocal groupings of dots and lines and by the use of sentences with double meanings. The accretive type of learning may be measured and scaled by using an increasingly elaborated maze with many turns. While such tests already form a part of many tests, they have not been differentiated and graded to bring out levels of intelligence as revealed in each type of learning.

BRIEF ACCOUNT OF REPRESENTATIVE TESTS OF NON-INTELLIGENT TRAITS

TRAIT	AUTHOR AND NAME OF TEST	BRIEF DESCRIPTION	TECHNIQUE EMPLOYED	PUBLISHED
1. Character	Woodworth's Personal Data Sheet	Test for evaluating the assets, liabilities, and compensations of the individual; list of traits about the person to be answered by those acquainted with him	True-false method	C. H. Stoelting and Co.
2. Emotional conditioning	Watson's Survey of Public Opinion	Tests tenacity and variety	Scale method	Union Theological Seminary, New York City
3. Expansion, ascendance, submission	Allport's Active-Passive Study	Subject reacts to a social situation as represented on paper: "You desire to board a boat or train to see a friend off. You feel it is important to do this; but the guard forbids you on obviously unnecessary technicalities. Do you obey silently, argue, or bluff your way past?"	Multiple-choice method	Unpublished

BRIEF ACCOUNT OF REPRESENTATIVE TESTS OF NON-INTELLIGENCE TRAITS (*Continued*)

TRAIT	AUTHOR AND NAME OF TEST	BRIEF DESCRIPTION	TECHNIQUE EMPLOYED	PUBLISHED
4. Emotionality	Pressey's X-0 Tests for Investigating the Emotions	Tests consist of list of words of highly emotional content and a list of crimes from mere misdemeanors to crimes of great enormity; standardized tests with tentative norms; adult form and juvenile form	Cross-out method; subject crosses out certain words, phrases, or pictures	C. H. Stoelting and Co.
5. Will power	Fernald's Achievement Capacity Test	Measures subject's will power in the persevering to stand on his toes as long as possible; low score considered as being due to unwillingness to suffer discomfort	Performance method	C. H. Stoelting and Co.
6. Trustworthiness	Voelker's Tracing and Opposites Test	Purpose is to find out if subject will cheat in an examination; paraffin paper is used to detect evidence of misconduct	Performance method	*Teachers College Contributions to Education*, No. 112, 1921
7. Moral judgment	Pressey's Mental Survey	Subject faced with suppositious dilemma: "If a man owes you $5 and refuses to pay, are you justified in keeping the $5 bill he happens to drop?"	Cross-out method	Indiana University

8. Emotional attitude	Myerson's Personality Tests	List of partial sentences with blanks to fill in; cards bearing several possible completions go with them: "A man who lives a pure life" (*a*) "will miss a lot of fun, (*b*) will gain the respect of all, (*c*) will be cheated by rogues."	Completion method	*Archives of Neurology and Psychiatry*, Vol. 1, 1919
9. Moral judgment	Cady's Moral Judgment Test	Pen-and-paper test of moral judgment based on 46 words: "Childish [people] are (1) very greatly to blame, (2) a good deal to blame, (3) a little to blame, (4) not at all to blame"; one of the four choices must be checked	Multiple-choice method	*Journal of Delinquency*, Monograph No. 2, 1923
10. Reliability	Voelker's Reliability Test	Subject asked to deliver a letter on time; evidence sought whether he read it since unsealed	Performance method	*Teachers College Contributions to Education*
11. Honesty	Voelker's Honesty Test	Subject sent to buy something; is given 25¢; examiner arranges with storekeeper to give overchange; honesty shown if subject returns the money	Performance method	*Teachers College Contributions to Education*

BRIEF ACCOUNT OF INVESTIGATIONS IN NON-INTELLIGENCE TRAITS

TRAIT OR TRAITS	NAME OF INVESTIGATION OR STUDY	BRIEF DESCRIPTION	NATURE OF RESPONSES MADE BY SUBJECT
1. Character traits	"An Experimental Study of Character Traits," R. O. Filter, *Journal of Applied Psychology*, Vol. 5, pp. 297–317, 1921	Speed of decision through two tests of self-assurance: (*a*) diagram shown and subject asked to reproduce on fingers by means of a string; (*b*) paper folding test (after Binet), asked whether he can draw the diagram when paper is unfolded, then draws it; (*c*) Downey Speed Decision Test; (*d*) Muensterburg-Gibson Test; (*e*) speed in evaluating actions; (*f*) speed in choosing between two pairs of items; (*g*) speed in comparing areas; (*h*) speed in deciding between two pairs of items — which is the greater evil?	Performs some task; chooses quickly
2. Social attitudes and interests	"A Test of Social Attitudes and Interest," H. Hart, *University of Iowa Studies in Child Welfare*, Vol. 2, No. 4, 1923	Consists of 150 stimulus words listed as follows: (*a*) things persons like or dislike; (*b*) things to read or study; (*c*) things one favors or disfavors	Subject manifests attitude by responses, which are scored as positive, neutral, or negative

3. Social perception	"An Experimental Study of the Growth of Social Perception," G. S. Gates, *Journal of Educational Psychology*, Vol. 8, pp. 449–463, 1923	Six photographs depicting (*a*) pain, (*b*) anger, (*c*) joy, (*d*) horror, (*e*) scorn, and (*f*) surprise	Subject labels the emotion portrayed
4. Personality	"Personality Test Involving the Principle of Multiple Choice," A. Myerson, *Archives of Neurology and Psychiatry*, Vol. 1, 1919	Test in the form of a story with certain parts left blank; nature of completion revelatory of character	Chooses parts that fit into deleted portions of story
5. Suggestibility (resistance to suggestion)	"A Study in Suggestibility in Children," M. Otis, *Archives of Psychology*, No. 70, 1924	(1) Directions Test: 40 items (20 suggestions and 20 blinds) arranged in order of difficulty (scale) (2) Suggestion Items: picture and leading questions asked about it; in form of scale	Follows directions and leading questions
6. Emotionality and expressiveness	"Criteria of Emotionality," C. Landis, R. Gallette, and C. Jacobson, *The Pedagogical Seminary*, June, 1925	Subjects: 12 men, 12 women, and a boy (*a*) Expressiveness (objective, subjective, and experimental); (*b*) stability (self-estimate of time consumed in recovering from emotional disturbance); (*c*) amount of	Answers questions; follows directions; submits to examination

BRIEF ACCOUNT OF INVESTIGATIONS IN NON-INTELLIGENCE TRAITS (*Continued*)

TRAIT OR TRAITS	NAME OF INVESTIGATION OR STUDY	BRIEF DESCRIPTION	NATURE OF RESPONSES MADE BY SUBJECT
		laughter; (*d*) emotionality and stability (rating scales); (*e*) standardized tests (Woodworth Questionnaire, Pressey X-O Test); (*f*) reaction (200 choice reactions); (*g*) height-weight ratio; (*h*) vocabulary test; (*i*) tapping test (3 ten-second periods before and after experiment); (*j*) fatigue coefficients; (*k*) systolic blood pressure (range and variability) Concludes that best criteria of emotionality are (*a*) increased reaction time as result of emotional upset; (*b*) a rating scale estimate; (*c*) the fatigue-index of tapping (Woodworth Questionnaire)	

Exercises for Self-Testing and Investigation

I. Rossanoff classifies personalities as
   (*a*) antisocial;
   (*b*) cyclothymic (behaving like a swinging pendulum);
   (*c*) shut-in or autistic;
   (*d*) epileptic.
   What methods would you employ in order to classify individuals according to these classifications?

II. Tests of non-intellectual traits have been divided into four classes:
   (*a*) those measuring ethical, moral, social, and religious traits;
   (*b*) those measuring character and personality traits;
   (*c*) those measuring interests and attitudes;
   (*d*) those measuring instincts and emotions.

   Indicate to which class or classes the following tests belong:
   1. Cady's trustworthiness test
   2. Kohs's ethical discrimination test
   3. Voelker's character tests
   4. Cady's incorrigibility test
   5. May's conscientiousness test
   6. May and Hartshorne's honesty test
   7. Watson's fair-mindedness test
   8. Voelker's truthfulness test
   9. Raubenheimer's book checking and overstatement tests
   10. Knight's book titles test

III. In the following statements of practice, procedures, and conditions one is the most acceptable, though the others enjoy a high degree of validity. Indicate by (*a*), (*b*), or (*c*) the most plausible one.
   1. Tests of character traits may be validated by
      (*a*) correlation between intelligence scores and character test scores.
      (*b*) correlations between the results of a test and the case histories and analyses.
      (*c*) correlation with teachers' ratings.

2. The most reliable methods of validation are by
   (*a*) correlation with ratings.
   (*b*) correlation with other objective tests.
   (*c*) age-gradation.
3. In addition to the I.Q., the most important additional predictive datum is furnished by
   (*a*) educational age.
   (*b*) degree of industriousness.
   (*c*) quality of secondary memory.
4. Character traits are best revealed by
   (*a*) observational records.
   (*b*) character tests.
   (*c*) teachers' estimates.
5. Industriousness is best developed by
   (*a*) exhortations to exert will power.
   (*b*) working with dominant interests.
   (*c*) resorting to emulation.
6. Politeness is a function of
   (*a*) character.
   (*b*) intelligence.
   (*c*) temperament.
7. Observational records should be maintained for periods of
   (*a*) one month.
   (*b*) one term.
   (*c*) one year.
8. All tests should be given
   (*a*) during the morning session.
   (*b*) during the afternoon session.
   (*c*) at any time.
9. Most character tests have
   (*a*) raw scores.
   (*b*) age levels.
   (*c*) character quotients.

10. The most desirable quality of personality is
    (*a*) consistency of action.
    (*b*) possession of ideals.
    (*c*) conformity to social criteria.

Bibliography

Badanes, J. and S. — *Observational Record*, published by the authors, Public School No. 173, Brooklyn, N. Y., 1920

Cady, V. M. — "The Psychology and Pathology of Personality," *Journal of Delinquency*, Vol. 7, 1922

Collins, M. — "Character and Temperament Tests," *British Journal of Psychology*, Vol. 16, 1921

Downey, J. E. — *Will-Temperament and Its Testing*, World Book, 1923

Fernald, G. G. — "The Defective Delinquent Class: Differentiating Tests," *American Journal of Insanity*, Vol. 68, 1912

Jung, C. G. — *Psychological Types*, Harcourt, Brace, 1923

Leaming, R. E. — "A Study of a Small Group of Irish-American Children," *Psychological Clinic*, Vol. 15, 1923

Lentz, T. F. — *An Experimental Method for the Discovery and Development of Tests of Character*," Teachers College, Columbia University, No. 180, 1925

May, M. A. — "The Present Status of the Will-Temperament Tests," *Journal of Applied Psychology*, Vol. 9, p. 50, 1925

May, M. A., and Hartshorne, H. — "First Steps toward a Scale for Measuring Attitudes," *Journal of Educational Psychology*, Vol. 17, 1926

Moore, H. T., and Gilliland, A. R. — "The Measurement of Aggressiveness," *Journal of Applied Psychology*, Vol. 5, 1921

Roback, A. A. — *The Psychology of Character*, Harcourt Brace, London, 1927

Smith, W. W. — *The Measurement of Emotion*, Harcourt Brace, London, 1923

Terman, L. M. — *Genetic Study of Genius*, Stanford University, 1925

Travis, R. C. — "The Measurement of Fundamental Character Traits," *Journal of Abnormal and Sociological Psychology*, Vol. 19, 1925

CHAPTER VIII

THE MENTAL DEFECTIVE

A. Historical Résumé

The mental testing movement arose out of the need for a segregative measure of extreme abnormality. Law and medicine and religion coöperated to establish norms for the measurement of defect. To Blackstone, "an idiot, or natural born fool, is one who hath no understanding from his nativity, and is therefore by law presumed never likely to attain any." Dr. Willes in 1678 declared that: "Idiocy and stupidity depend on a lack of judgment and intelligence, the actual rational mind being not concerned; the brain is the seat of sickness, which rests on a lack of imagination and memory, the seat of which is in the brain. The imagination is located in the *Corpus Callosum*, or the white substance; the memory in the cortical substance. So if imbecility and stupidity arise, the cause is in the brain region involved, or in the animal spirits, or in both." The ecclesiastical definition is best expressed by Heimroth (1818), who wrote: "Where Satan is there is weakness, darkness, hatred, and destruction everywhere. An evil spirit abides, therefore, in the mentally deranged; they are the truly possessed."

To-day defect is judged by various criteria: (1) the social-economic, (2) the medical, (3) the pedagogical, (4) the psychological, and (5) the biological. These criteria are not, however, mutually exclusive. The intimate relations that obtain among all sciences, both physical and social, are bound to lead to "cultural borrowings"; they differ mainly in the emphasis placed.

(1) *Social-economic Criterion*

Great Britain leans towards a social-economic interpretation of feeble-mindedness. The Mental Deficiency Act for 1913 establishes three grades: "The feeble-minded (termed *morons* in the United States) are persons in whose case there exists from birth or from an early age mental defectiveness not amounting to imbecility, yet so pronounced that they require care, supervision, and control for their own protection or for the protection of others, or, in the case of children, they by reason of such defectiveness appear to be permanently incapable of receiving proper benefit from the instruction in ordinary schools.

"*Imbeciles* are persons in whose case there exists from birth or from an early age mental defectiveness not amounting to idiocy, yet so pronounced that they are incapable of managing themselves or their affairs, or, in the case of children, of being taught to do so.

"*Idiots* are persons so deeply defective in mind from birth or from an early age as to be unable to guard themselves against common physical dangers."

(2) *Medical Criterion*

There is marked divergence among medical authorities as to the nature of defect. We shall cite two representative definitions which do not attempt a differential diagnosis. Bourneville defines idiocy as an "arrested development, either congenital or acquired, of the intellectual, moral, and emotional faculties, which may or may not be accompanied by motor difficulties and perversions of instinct." Dr. Tredgold's definition is eclectic: "The condition is a psychological one, although the criterion is a social one, and we may accordingly define amentia as a state of restricted potentiality for, or arrest of, cerebral development, in consequence of which the person affected is incapable at maturity of so adapting himself to his environment or to the requirements of the community as to maintain existence independently or without external support."

(3) *Pedagogical Criterion*

This criterion is based upon the degree of pedagogic retardation and is applicable only to feeble-mindedness, which is defined as a condition of retardation amounting to three or more years. The child who is three years behind in his class is presumed to be defective.

(4) *Psychological Criterion*

This is a quantitative measure in terms of the intelligence quotient. It recognizes three grades of defect: normal, I.Q. 40 to 70; imbecile, I.Q. 20 to 40; idiot,

I.Q. 20 to 0. Imbeciles and morons are subdivided into low-grade, mid-grade, and half-grade.

(5) *Biological Criterion*

The biological criterion is one of adequacy of adjustment. It is comprehensive enough to include all forms of abnormal behavior penalized by society. It is social to the extent of stressing the lack of prudential intelligence. It defers to medical lore in its insistence on the presence at times of physiological derangement (nervous and glandular). It borrows from psychology the instruments designed for mental testing, and it goes to the schoolroom for corroborative pedagogical evidence. It adds, however, the causative factor of maladjustment as evidenced by the perverting influence of complexes. The biological criterion does not base itself solely on bodily conditions, on pedagogical insufficiency, on mentalistic trends, or on social maladjustment, the reason being that the human animal is an integrative organism, whose normal functioning is conditioned by personality trends, both innate and acquired, coördinated into a complete personality. Insufficiency or perversion in any major personality trait is bound to react on the other traits, and the normal functioning of one trait involves the coöperation, direct or indirect, of the rest.

The study of the total personality gives to the term "insanity" a less opprobrious meaning. Difficulties of adjustment beset all individuals. Only the normal

attain stability without serious impairment of behavior. As Dr. McDonald forcibly puts it, " Substantially every individual at some time or other during his life is exposed, in many cases repeatedly, to many of the so-called exciting causes of insanity, both mental and physical."

The biological position may be stated thus: Children may be classified as subnormal and normal. An individual may be called subnormal who by reason of innate insufficiency, physical, mental, and emotional, or by reason of an acquired pathological condition of body or mind, or through perversion of the emotional life, or by reason of all three causes, cannot maintain himself adequately in school. There are degrees of subnormality, but the concern of the school is with three large types: (1) the subnormal in intelligence, (2) those subnormal (neurotic) in temperament and character (moral defectives and social delinquents), and (3) those who partake of the defects enumerated in (1) and (2). For the sake of pedagogical simplicity we shall limit our discussion to the mental defective and the neurotic.

B. Characteristics of the Mental Defective

(1) *Specific Characteristics*

1. Bodily Concomitants of Mental Defect. Since the total personality is implicated in mental life, it behooves us to examine the bodily conditions that may accompany feeble-mindedness. Statistical information available to date indicates that the feeble-minded are shorter

and lighter, age for age, than the normal; that, Lembroso to the contrary, single deformities (stigmata) abound in equal degree in the normal and subnormal; that the feeble-minded are inferior in motor ability (grip, coördination, and speed of voluntary attention); that their gait is awkward and shambling; and that malformations occur in the case of those of the lowest level of feeble-mindedness (hydrocephaly, microcephaly, Mongolianism, and Gigantism). It has also been established that this class does display anomalies of structure, but that diagnosis of mental deficiency cannot be made on this basis.

2. The Instinctive and Emotional Concomitants. We have seen that instincts and emotions are, in early life, determinants of conduct. Later, conditioning, education, and trial and error modify and elaborate the inborn instincts and emotions. They become associated into a motivating body called the unconscious. This in turn is checked, directed, rationalized, and socialized by the application of intelligence. The defective, being of feeble mentality, interposes no restraining power between desire and gratification. His behavior, then, is likely to be masked by a disconcerting directness. Since he is unburdened with inhibitions, seeing is taking and desiring is gratifying in the absence of physical restraints. The emotional life, however, is not markedly different from that of normal individuals, defectiveness being traceable to the low level of intelligence. The primary feeling trends are

just as powerful, but the emotional reactions remain primitive, not having become intellectualized through sublimation. The defective has not learned to attach the fear and love responses (and their variants) to situations that do not normally call them out. His sense of humor, for instance, has little of the subtle in it.

3. Mental Characteristics of Feeble-Mindedness. Defect seems to grow acute above the level of perception. The senses of the feeble-minded are on a parity with those of normal individuals. They are successful in building up percepts; but these remain, in many instances, at a level little higher than mere sensations. Moreover, they are disorganized and without meaning. As Hollingworth succinctly puts it, " He [the defective] fails to learn to read, not because his retina is insensitive to black and white, but because he lacks the intellectual development which is prerequisite to the interpretation of the black marks which he sees." This lack of interpretation is most marked in dealing with abstract symbols. Concrete phenomena elicit a keener understanding, as witness his astonishing proficiency in the manual arts.

He also differs sensibly from the normal in the ability to adapt himself to new situations, especially when the new situation is compounded of many elements. This is complicated by an uncritical attitude as infantile as it is credulous. Norsworthy evaluated these specific

defects on the basis of mental tests. She found that feeble-minded people are two and a half times as far from the median for children in general in tests like the opposites test or the genus-species test as they are in tests like the "A" test or the perception of weight. The feeble-minded child may be weak on all sides of his mental make-up, in the sense that he is not of normal ability in any respect, but this is not telling the whole story. From the point of view of the psychologist and the educator it is fully as important to know that the ament's perceptive powers are almost two and a half times as strong and accurate as his intellectual powers, and almost half as strong as his powers of memory, as it is to know that he is weaker than an ordinary child in all these particulars!

(2) *Summary of Characteristics*

We may present these characteristics by combining in one table the mental, physical, and emotional traits which may or may not coexist in the same individual. While we no longer consider the presence of physical abnormalities evidence of degeneration — what the followers of Lombroso called "stigmata of degeneration" — it behooves us, nevertheless, to give heed to them. We must, however, stress mental and emotional aberrancies as possible determinants of feeble-mindedness.

	Quality	Defective
Mental traits	Reception	Weak, devoid of subtle meaning
	Memory	Poor retentive powers
	Reasoning	Normal in concrete situations; poor in abstract thinking
	Self-criticism	Suggestible, uncritical, infantile
	Sense equipment	Suggestible, normal in acuity
Physical traits	Weight and height	Inferior
	Stigmata	Frequent, particularly malformations of skully palate, teeth, skeleton, and eyes
	Glandular disturbance	Cretinism, Gigantism
	Nervous disorders	Chorea, tic, St. Vitus Dance, epilepsy
	Gait	Shambling, awkward
	Motor ability	Poor coördination, weak grip, poor speed, involuntary movements
	Vital capacity	Poor lung expansion as measured by spirometer
	General system	Remains infantile or organs of reproduction malformed
Instincts and emotional life	Instincts	Primitive in organization; readily activated; productive of delinquencies, such as stealing, assault, truancy
	Emotions	Uninhibited, unintellectualized, unsocialized

C. Testing and Establishing Mental Deficiency

The cause of feeble-mindedness being hereditary or due to injury or disease of the nervous system, we must study the individual from the point of view of predisposing circumstances as well as present status. The

former may be established through (1) a study of the family and individual history and (2) medical examination. The prevailing practice is to include both under the heading "medical." The examination is conducted under the following headings:

I. History of the family to establish hereditary factors
   (1) Mental characteristics of father, mother, and grandparents
   (2) Diseases to which family seems predisposed

II. History of the individual to reveal dominant mental and emotional characteristics, such as suggestibility, brightness, feelings of inefficiency, passivity, etc.
   (1) Early childhood
   When did he learn to walk? to talk?
   What diseases did he have?
   (2) Injuries and diseases in later life

III. Physical examination
   (1) Nutrition
   (2) Bones and joints
   (3) Circulatory system
   (4) Respiratory system
   (5) Gastro-intestinal organs
   (6) Condition of blood
   (7) Glands

IV. Neurological examination
   (1) Condition of sense organs

(2) Movement — rapidity and accuracy of movement; gait
(3) Eyes open and closed
(4) Reflexes: patellar, contralateral, abductor, Achilles tendon, Babinski

V. Psychological examination

VI. Pedagogical examination

A word about the psychological examination. The present trend is decidedly against placing too great a dependence upon an exclusively psychological test. The reasons adduced are many: First, the Stanford-Binet Scale (and this is the preferred measuring instrument) is not sufficiently discriminating in borderline cases or in cases complicated by pedagogical retardation. The upper limen (limit) of feeble-mindedness is but an arbitrary figure. As a matter of fact there is no "dead-line" of normality; there are too many recognizable intermediate stages which, at a later stage, permit passing the supreme test of normality, what Huey calls "the ability to float in society." Secondly, Binet himself felt constrained to utter this warning: "In spite of the system of annotation which we have devised, we think it is the duty of the experimenter to *judge, weigh,* and *examine* the replies. Our method is not an automatic weighing machine like those in railway stations, which register automatically the weight of a person, without intervention or assistance." Again he remarks that the limit which "separates debility

from the normal state" is "not fixed but variable according to circumstances." Thirdly, the intelligence quotient is not accurately and infallibly diagnostic of feeble-mindedness. Wallin * cites three illuminating cases to illustrate the poor prognosticability of the I.Q.: "M., age 16.83, B.S. age (mental age on Binet Scale), 8.4, has a quotient of .49, and should be an imbecile according to one proposed standard. He is a moron. M., age 9.66, B.S. age 6.4, has a quotient of .66, and should be a moron. . . . He was diagnosed as an imbecile and his subsequent record indicates the correctness of the diagnosis. M., age 14.08, B.S. age 6.4, has a lower quotient, .73, than M., age 8.58, B.S. age 6.4, quotient .74, but the former has proved after three years of observation to be only backward, while the latter still grades as an imbecile, although he may reach the status of a moron." Fourth, the best that the mental test can do is to serve as an aid in diagnosis, just as many medical tests are used in medical diagnosis. Its value lies in indicating a line of action and in corroborating other symptoms. But even as a first line of attack, its reliability inheres less in the quantitative delimitation than in the interpretative value of the specific answers. A trained psychologist learns more from the individual's mode of answering than from the answers themselves.

The inadequacy of mental testing as the sole criterion

* From *Problems of Subnormality*, copyright, 1917, by World Book Company, publishers, Yonkers-on-Hudson, N. Y.

of mental sufficiency is more evident when applied to potential feeble-mindedness than to grades of normality. The cause is not far to seek and may be summarized as follows:

1. We are attempting to test innate capacity apart from education and training. Since defect is innate, we must devise tests of original endowment.

2. We must abandon the idea of measuring intelligence by means of one test, no matter how comprehensive.

3. We are prone to place too much faith in average measures; no norm is precise enough to allow for individual variability, which enters as a disturbing element even within classic types.

4. We are trying to measure how much general intelligence the individual has, forgetting that for purpose of education the fact of greatest pertinency is how much of it he possesses in any given direction. We are measuring by levels instead of indicating both peaks and depressions.

5. We are replacing scientific insight with automatic tests. The giving of a test is the least important phase of mental testing. Psychologists might paraphrase a famous dictum thus: "We don't care who gives the test as long as we interpret it!"

D. Teaching the Mental Defective

(1) *Historical Survey of Methods Employed in Teaching Defectives*

To Jean Marie Itard (1775–1838) must be accorded the distinction of having addressed himself to the task with great competence. Itard took exception to Pinel's statement that the " savage of Aveyron " was an idiot and therefore beyond pedagogic help. Accordingly he undertook his education, setting himself the following program:

1. To develop his general intelligence. He tried to get the boy to comprehend the meaning by associating words with most elementary needs. He also made use of all spontaneous tendencies as they manifested themselves, omitting no opportunity to tie up instinctive trends with useful acts.

2. To socialize his conduct by making civilized life more pleasant than primitive existence in the woods. Tact, patience, and forbearance characterized their relations.

3. To enlarge the content of his mind.

4. To refine his senses. He attempted to train all the senses by any means short of harmful exposure. He began with the sense of touch. He gave him baths of varying temperatures to develop the sense of cold and warmth. He followed this up by training the senses of taste and smell, offering stimuli that might compete with the less refined one the boy greatly favored.

Mention might be made of Itard's predecessor, Pereira, who achieved a measure of success in educating deaf-mutes. Pereira enunciated a great pedagogic truth when he advanced the theory that sense training must precede intellectual training, and that when one of the senses is atrophied, another one might be trained to take its place. The blind may be taught to substitute for visual perception touch perception (Braille). This training may proceed to the point of greatest acuity, so that the favored sense may compensate for the defective one. Montessori has seized upon this idea and has made it fundamental in her scheme of education. Thus Montessori provides her pupils with a great store of tactual and kinæsthetic sensations of shapes of letters and figures before they are permitted to handle paper and pencil. The advantage of this preliminary training is that information is built up without conscious effort. Memory is enhanced because retention seems to be greatest when mediated by the sense of touch.

We need not detail here Rousseau's contributions except to mention that he learned a great deal about sense training through Pereira, who was a personal friend. Seguin (1812–1880) with his so-called physiological method improved upon his predecessors. He stressed bodily vigor and sense training and muscular coördination, claiming, and rightly so, that unless the idiot can see the outer world correctly, all verbal education is worse than useless. He therefore devised apparatus for energizing the muscles and for producing

all fundamental movements requisite in learning how to stand and walk. He resorted to "treading the floor of a treadmill." Motivation was not neglected. When the child had hurt his hands trying to learn the finer points of locomotion, he was not left long in saddened contemplation of his condition; but directing him to extend his hands, Seguin put on each a bright apple. "He, partly to feel the coolness on all the burning surfaces, partly to let the apple fall, will contract his fingers and get a circular, equable, willed prehension of them."

But the stimulation of the senses was not merely mechanical. He sought to replenish the energy sources behind the functioning of the sense organ. He expressed this principle in philosophic language. "Each sense," he warns us, "must be taught as a function, and taught besides as a faculty." What he meant to say is that, where sensory deficiency is a matter of weak will, the will must be strengthened through the multiplication of desires, on the assumption of proven validity that, where there is specific impulse, expression will follow. The same procedure was followed in training the other senses. In training how to speak, the mechanism of speech is taught first; then a desire is created to use it.

Seguin's contribution to the method may be summarized thus:

1. Education must proceed from physical adjustment to intellectual competence.

2. The senses must be trained through the handling of material; muscular coördination must be ingrained, through imitation or spontaneous manipulation of stimuli if possible, through mechanical means if necessary.

3. Physical, sensory, and intellectual progress must be simultaneous.

4. Instruction must wait upon the need for it; and, as a corollary to it, the desire must be inculcated.

We have already mentioned Montessori. Seguin appealed to desire and motive, but like Rousseau he became lost in theories. Montessori is nothing if not practical. She therefore set out to devise apparatus through the manipulation of which desire is created. The apparatus is carefully graded so that the mastery of one piece motivates the mastery of the more elaborate mechanism, gradually leading, by insensible gradations, to the acquisition of educational processes. Great ingenuity characterizes the construction of the apparatus, each part fitting into the only space available, thereby affording opportunities for the exercise of self-instruction and auto-criticism. She teaches reading and writing in rough games which are so much like games that learning to read and write lose the solemnity with which the school invests them. Moreover, reading and writing are acquired through an unstudied acquaintance with their sensory background extended to many senses. Not only is there multiple sense appeal, but there is intimate sense appeal, many of the senses making a great many contacts with the sense material.

Thus letters are presented to the touch in various guises, rough, smooth, wide and narrow, large and small.

Her devotion to sense training has led Montessori astray. Seguin saw in each sense an extension of the mind; to him a sense was both an instrument and a faculty. Montessori would train the senses for the sake of training them. It is now known that, unless the individual is destined for a vocation requiring sensory acuity of a superior order, the refinement of the senses serves no intellectual purpose. A person may develop a keen sight, a delicate sense of touch, and acute hearing without adding one cubit to his intellectual stature. This shortcoming need not detract from the value of the Montessori method. She has made it possible for the defective to attain some measure of social and intellectual competence.

(2) *Modern Methods of Educating Defectives*

Modern psychology would temper the pedagogical winds to the protective equipment of the defective. Experimental studies have established the following pedagogical weaknesses of the feeble-minded:

1. The feeble-minded do not learn as rapidly as the normal.

2. They do not improve as much or as rapidly as the normal when subjected to the same kind of training.

3. They do not transfer their training in any one task to other tasks of similar degree of difficulty and range. Every new situation must be presented anew;

they are unable to combine similar elements of old situations to meet the requirements of slightly altered circumstances.

4. The mental age of the defective child erects its own educational barriers; a child can learn no more than its mental age permits. Experimental data are still lacking to establish definitely the kind and amount of learning possible. But it is fairly certain that certain subjects, and phases of others, are forever sealed information to those failing to attain a definite mental age. Thus it may be stated with some degree of certainty that the feeble-minded can never learn fractions or syntax.

5. The feeble-minded can learn just as readily as the normal of corresponding mental age; only the former shows a greater irregularity in the curve of learning.

6. The innate capacity to learn cannot be increased.

(3) *Summary of Methods*

It is reasonable to conclude, therefore, that defectives must be taught differently; that the school must emphasize the inculcation of social habits, each one of which must be ingrained separately; that education should proceed through the senses, particularly the motor senses; that concrete work should predominate; that only that degree of abstract learning should be required which accords with the mental age; that all efforts should be exerted to render the individual self-sustaining economically; and that morality must be

taught in terms of concrete habits rather than through the inculcation of ideals.

Exercises for Self-Testing and Investigation *

I. 1. Accomplishment in subject matter and chronological age may serve as a basis for grade placement if supplemented by

(*a*) I.Q.
(*b*) pooled teachers' judgment.
(*c*) mental age.

2. Children with a mental age below six generally do

(*a*) poor work.
(*b*) average work.
(*c*) satisfactory work.

3. Pupils with low mental rating do best in

(*a*) academic work.
(*b*) manual work.
(*c*) commercial work.

4. The I.Q.

(*a*) remains relatively constant.
(*b*) remains inconstant at higher ages.
(*c*) varies with change in social environment.

5. In grading pupils, classes should be formed of

(*a*) gifted, bright, normal, and slow.
(*b*) bright, normal, and slow.
(*c*) bright over-age, bright normal-age, and bright under-age;
normal over-age, normal normal-age, and normal under-age;
slow over-age, slow normal-age, and slow under-age.

* Planned for use with Chapters IX and X also.

6. To discover nascent powers, children should be tested
   (*a*) upon entering school, at the end of the sixth grade, and at the end of the eighth grade.
   (*b*) in kindergarten, 4B, and 8B.
   (*c*) in kindergarten, 4B, 7A, and 8B.
7. The chief cause of failure among school children is
   (*a*) late entrance into school.
   (*b*) truancy.
   (*c*) poor mentality.
8. A curriculum is best adjusted to the individuality of the child by
   (*a*) insisting on minimal essentials.
   (*b*) consulting the child's interests.
   (*c*) adapting the traditional curriculum to the child's level of intelligence.
9. The most neglected class of children in our schools is
   (*a*) the slow pupil.
   (*b*) the feeble-minded.
   (*c*) the gifted.
10. Segregation of gifted pupils into special groups
   (*a*) enhances emulation.
   (*b*) stifles emulation.
   (*c*) promotes egotism.

II. Indicate which of the following statements are true and which are false:

1. A backward pupil who is sincere is apt to be overrated.
2. A verbalist may possess intelligence of a high order if his utterances possess coherence.
3. It is doubtful whether there exists such a defect as spelling disability.
4. The gifted are taller than unselected children.
5. The size of the body is negatively correlated with intellect.
6. Gifted children are ineffective in athletic games.

7. Gifted children rank higher in character than those of lesser endowment.
8. Gifted children are not inquisitive.
9. Nervous instability is a sign of feeble-mindedness.
10. Social intelligence is a happy blend of temperamental, physical, and intellectual traits.
11. Gifted girls are not amenable to traditional girls' games.
12. Gifted children are interested in stock juvenile stories.
13. Gifted children are difficult to discipline.
14. Gifted children are receptive to appeals for respect of the rights of others.
15. Dull pupils learn spelling through sheer repetition.
16. Dull pupils usually become good stenographers.
17. Efficiency in handwriting correlates highly with intelligence.
18. Backward pupils lack self-criticism.
19. Backward children do well in technical grammar.
20. For dull pupils the teaching of phonetics adds a confusing element.
21. Dull pupils read for content rather than for word calling.
22. Bright pupils manifest an overwhelming superiority in language.
23. Neurotics are usually of mediocre intelligence.
24. Dull pupils do poorly in adding long columns.
25. Dull pupils do well in arithmetical problems.*

* The Bibliography for Chapters VIII and IX will be found on pages 324–326.

CHAPTER IX

THE NEUROTIC

A. Difference between the Defective and the Neurotic

One indication of the chaotic nature of our present knowledge of abnormal psychology is found in the multiplicity of terms evidencing great duplication and even contradictoriness. We have seen that psychiatry and psychology employ a disparate terminology varying with their nationalities. Equally unsatisfactory is the classification of neurotics. They are variously known as psychopathic, neuropathic, and unstable. In popular parlance they are called "nervous." In school they comprise that large class designated euphemistically as "difficult" or "problem cases." From this class are recruited the truants and the delinquents.

Where shall we place them with regard to the normal children? They are undoubtedly subnormal or deviates. But their subnormality is not a matter of inferior intelligence level. Herein they differ from the defective. The defective faces the world with a full assortment of intellectual coins, but they are for the most part of low denomination. He can buy but a limited

quantity of the world's goods. The neurotic comes supplied with a full complement of intellectual coinage, but his purchases are characterized either by extreme niggardliness or extravagance.

Psychologically, the defective suffers from a low level of mentality; the neurotic, on the contrary, exhibits a poor quality of intelligence. The former is quantitative; the latter is qualitative. Furthermore, the neurotic lacks control of his emotional life, which is more complicated than that of the defective who, because of the restricted view of his environment, is singularly free from conflicts and suppressions. The neurotic is vouchsafed the same emotional experiences as is afforded the normal child, but owing to a poorly functioning mentality is apt to give them an abnormal bent.

B. Characteristics of the Psychopathic Individual

Characteristics range from minute anomalies and idiosyncracies to extreme dementia. The neuropathic may deteriorate into a psychopathic. Again, just as in the case of defectives we cannot definitely assign limits between normality and feeble-mindedness, so in differentiating the neuropathic from the normal it is difficult to speak with finality. When a normal person is fatigued, he may act abnormally. He may forget names, act absent-mindedly, display extreme nervousness, and yield to fears and rages. But this functional

deviation is temporary. On the other hand, the neurotic displays abnormal symptoms unmitigated by fatigue.

We may enumerate certain traits of the neurotic constitution. They are verbalism or excessive talking, usually more plausible than rational; mutism, the opposite of verbalism; irrational anger; automatic habits; overinquisitiveness; inability to get along with one's fellows; solitariness; pathological lying; impatience with all forms of constraint; delinquent conduct; amoral and immoral behavior; "peculiarities"; "mannerisms"; and "oddities." This by no means exhausts the list, but it is representative of the manner of reacting.

C. Causes of Psychopathy

There may be physical causes, mental causes, or a combination of these two. Here, too, heredity plays a predominant rôle either as a direct taint (syphilis) or as a predisposing cause. The physical causes are many, though their exact influence is still undetermined. Psychiatrists are inclined to regard psychopathy as "functional," that is, arising "without apparent organic lesions or loss." However, there is general agreement with Ziehen that psychopathy involves "constitutional inferiority." There are psychiatrists who make the sweeping claim that, since there can be no function without structure, there can be no abnormal function without abnormal structure. They adduce the further argument that, since both the body and mind develop from

one and the same germ cell, any change in one is bound to affect the other. It is known that anatomical lesions in other organs of the body change the metabolism, contaminating the blood. The blood carries the abnormal products to the brain and disturbs its chemical and nutritional status. It is therefore argued that, since lesions react upon the brain, is it not possible that other organic disturbances of lesser abnormality may also involve the brain?

Mental factors alone may produce disordered mentality. Worry, shock, grief, mental overwork produce sleeplessness, fatigue, and nutritional disturbance. They lower the vitality of the organism and thereby cause latent infections to become active. These conditions also play havoc with the normal activity of the endocrine glands, thus aggravating the mental stress. In support of their contention, these psychiatrists cite many cases in which patients suffering from clouded mentalities have been restored to mental health by extracting infected or impacted teeth or by removing infected portions of the intestinal tract. Of 750 cases diagnosed as "neurotic" — "pests," doctors call them — 48% were found to have a correlative physical disorder, and of the remaining 52% the distinguished surgeon Dr. Mayo said in 1913: "They [the disorders] are all due to a physical cause. Ten years from now 60% will be so diagnosed; twenty-five years from now 70%, and a century or less from now we shall be able to find and remove the cause of 90%." Dr. Harry A.

Cotton, Director of the Psychiatric Clinic for Correctional Institutions of New Jersey, justifies Dr. Mayo's prophecy with a remarkable collection of cases cured by operative methods.

On the other hand, we must give a hearing to the proponents of functional causes of psychopathy, the Freudians. They see in wish fulfillment and conflict the cause of many mental and emotional disturbances. We shall examine their position in some detail.

(1) *The Freudian "Complex" and "Neurosis"*

The Freudians divide personality into an unconscious and a conscious. Into the unconscious gravitate our wishes and desires born out of a struggle, more or less intense, between the instincts and the environment. They ascribe this struggle to a grudging recognition of the exigencies of reality. Rare indeed is the situation in which the individual finds no impediment to a free exercise of the instinctive and emotional tendencies. Most individuals find their desires at war with social taboos and react in one of two ways: they replace these troublesome wishes by others less objectionable; they capitulate to reality without bitterness and without repining. This is the normal adaptation. However, there is a large number of less fortunate persons who are loath to surrender their pet wishes because their fulfillment is in a consonance with their pleasure principle; they are chary of giving up or repressing any desire that holds out so much promise of inner satisfaction. When

reality grows exigent, when the environment in the person of parent, teacher, or employer rides roughshod over his protestations, recourse is had to artifices which preserve the shadow of victory. The individual pretends that his wish is still supreme, at least in his own inner world of dream and phantasy.

This dream life is not necessarily a sign of abnormality. Since very few of us can make a frictionless adjustment to reality, phantasying is part and parcel of our behavior. Our claim to normality rests on the fact that, when we build castles in the air, we know that the blue prints call for neither brick nor mortar. We indulge in this architectural pastime because we wish to flatter our ego. This is not the attitude of the maladjusted individual. He builds castles, moves into them, and has to be dispossessed forcibly by the sheriff, reality.

The escape from reality is effected through a series of mechanisms. *Conversion* is one of them. They are mostly of a hysterical nature. Children find that parental attention is at its intensest during a severe illness. They therefore simulate illness in order to evoke solicitous attention on the part of parents. Soon simulation is not necessary; they actually feel the pain. Phister cites the case of a young man of seventeen who developed a pain in the left arm. The pain made itself felt after his father's insistence to have him transferred from the school he was attending. Here is a more complicated form of conversion. The youth is faced

with an intolerable situation; he hates to leave school. He therefore resorts to an infantile ruse. He converts the disagreeable feeling into a physical pain. The variety of such conversions is astounding. The resistance to reality takes such serious forms as asthma, sore throat, stammering, faint attacks, hay fever, gastro-intestinal disturbances, and diabetes. They have no existence outside the patient's mind.

Rationalization is another protective mechanism. The ineffectual person is most facile in explaining, with great circumstance and subtle sophistry, the reason for his untoward conduct. The " sophisms of the indolent " are woven in a pretentious tapestry of logic. As Dr. White wittily remarks, " [The alcoholic] drinks because it is hot or because it is cold, because he cannot refuse a friend, or one drink won't hurt him, or he did not really mean to drink that evening at all."

Compensation is the mechanism that is set in motion to restore the emotional balance upset by an uncomplimentary awareness of inferiority. The world exposes shortcomings ruthlessly and unfeelingly. If a boy's inferiority in scholarship is made grossly patent, he may seek balm and compensation by playing truant, thus effecting an escape, or by consorting with the gang who have learned to respect his prowess.

Protective mechanisms are not always efficacious. These pseudo-adjustments to reality fail when the pleasure motive and the reality motive come into violent contact. The shock may dislodge the wish and send it

toppling into the unconscious. It does not remain there inert. It may continue to make feeble efforts to reinstate itself as a slyly motivating complex or it may effectually block the channel for outgoing motives and dominate the unconscious. The complex deteriorates into a neurosis, and the individual is taken captive by it.

These neuroses manifest themselves in ways that are unintelligible to the normal individual. To the psychiatrist the behavior is susceptible of explanation. Charcot, an eminent French physician, was attracted to hysterical subjects. He found that their condition was traceable to certain painful ideas that had dropped out of consciousness. His pupil Janet ascribed hysteria to a splitting up of personality into subpersonalities, each one of which was capable of initiating action independently of the dominant personality. This autonomous fractionalization of personality can be reincorporated into a unitary personality when the individual is placed in a hypnotic state, for hypnosis neutralizes repression and painful memories are then shorn of their objectionable features — memories are revived without their emotional background. Janet was wont to suggest to hypnotized patients that this experience which the patient had so calmly reinstated in consciousness was a most matter-of-fact experience, devoid of any great significance, and that it therefore should not occasion any distress. One of Janet's patients, Marie, suffered from intermittent blindness of the left eye. Hypnosis was

resorted to, and one day Marie recalled that when she was six years of age she was forced to sleep in the same bed with a child having a scrofulous affection of the left eye. This incident caused an intense emotional disturbance. It happened that Marie was similarly infected, and when the ailment disappeared it left two alarming symptoms which persisted, loss of sensation in the left side of the face and blindness in the left eye. At the conclusion of this recital Janet made the therapeutic suggestion that the enforced company of the sick girl was not so terrible an experience, and when Marie awakened the symptoms had entirely disappeared.

Breuer and Freud became convinced that suggestion as a motive was powerless to effect permanent cures. They argued that the patient must be made to revive the experience in its emotional setting, taking care that the emotion aroused was restored in its original intensity; that, following this restoration, the patient must be encouraged to review it rationally and calmly. Breuer and Freud differed from their predecessors in emphasizing the centrality of the patient's ego. He permitted himself to be subjugated by an emotionally charged idea and he himself must break the fetters. In other words, if in building one's personality one builds wrongly, one must be made to tear down and rebuild. But the psychoanalyst needs to point out the building violations. The patient is powerless to discover them unaided.

(2) *Summary*

It matters very much whether we espouse the medical or the psychoanalytic view of psychopathy. If the mental factors are secondary, we must hospitalize delinquents. If the " complex " is at the root of the evil, we should attempt to reëducate the sufferer so as to force all complexes into consciousness. But it is not necessary to make a choice. The truth lies somewhere between: body and mind are interdependent and react to each other in subtle ways — much too subtle to warrant dogmatic statements. It is quite possible that the riddle of personality lies hidden in the endocrine glands. Glandular susceptibility to external stimulation may be at the bottom of all purely mental causes of psychopathy. Again, the vast array of ideas thronging the most impoverished intellect may well cause acute disharmonies. Our finite mind cannot enlarge its horizons indefinitely. Civilization, with its myriads of new stimuli and new ideas, is fast exceeding the mind's innate flexibility. Isn't the pitcher, progress, putting ideas over a little bit too fast for the mind's batting eye? We compose questions much more readily than we can answer them, and the doubts that assail us must make us pause before taking a definite stand. Whatever the true causes of psychopathy, the symptoms and social implications are definable and ascertainable.

D. Means of Determining Psychopathy

The neurotic is a highly complicated individual and should therefore be subjected to more exhaustive tests than will suffice for the defective. Aside from greater complexity of character there is marked difference in variability. The neurotic is an extreme individualist; he resists all efforts to destroy his uniqueness. It is idle to attempt subsuming him under "averages" and to equate him with "norms." It is not an exaggeration to say that each one is a "case" unlike all other cases. Differential diagnosis, then, must seek out all peculiarly individualistic attributes which in the normal would hardly excite comment. What in the normal is a mere "slip" must be considered the very essence of the psychopathy. The examination in both its physical and psychological aspects must be exhaustive, thoroughgoing, and circumstantial. The psychologist's participation in the survey of a neurotic may take these forms:

1. Testing with the Stanford-Binet or short Binet tests. What is diagnostic of psychopathy is not mental age as the intellectual range above the basal year (the highest mental year in which the subject answers all the questions). The neurotic is apt to show unexpected blind spots alternating with brilliant answers at progressing levels. Doll's investigation affords grounds for inferring instability from a variability extending four years above the basal year. Another item of great diagnostic value is qualitative variability. The

weight of experimental data points to the differentiating character of some tests. Thus psychopaths test well in rote memory, comprehension, and the use of abstract terms. They fall down in rhymes, reaction time, weights, constructing sentences from three words, solving inclosed boxes, and reproducing the content of material read.

There is still another qualitative test. Of utmost importance is the nature of each response and its concomitant behavior. After subtracting from each response its standardized elements, there remains, in the case of the neurotic, a rich and varied residue of irrelevant and inconsequent material that must be subjected to critical analysis. There is, for instance, the half-coherent reply, the injection of the personal pronoun " I," the gratuitous interpretation of definitions, the startling explanation of the nature of the absurdities, the monosyllabic answer where fullness of response is called for, and the repetition of the same word or a few words to fill the sixty-word quota. Again, the examinee may display peculiar mannerisms — blushing, twitching of facial muscles, stammering, getting up, roaming about the room, or displaying unaccountable excitement when certain tests are given.

2. The psychologists may throw light upon the insidious complexes behind the psychopathy. The standard test is that of Kent and Rossanoff. In this connection the suggestion of Frank Watts * is worthy of

* *Abnormal Psychology and Education*, Appleton, 1924.

adoption, especially when we consider the inconclusiveness of any word association tests found in this country. He advises that the psychologist get a line on the neurotic's trouble and construct a word association test embodying all the emotionalized ideas. He cites an instance, of common enough occurrence, which involved the use of such a test.

As a result of a theft committed in a certain class it became necessary to find the culprit. The article stolen was a brown leather purse which had been left by the teacher on the top of her desk alongside an upturned waste-paper basket. The culprit had upset a pot of red ink in his haste. By a process of elimination two boys, A and B, were suspected. They were invited to take the following word association test. The italicized words were intended to serve as complex indicators, the probers for the suppressed thoughts occasioned by the breach of social decorum. The words were read twice and time reaction noted.

Culprit A betrayed himself by the following signs:

Length of reaction time 3 to 5 seconds for each damaging reaction

Evidence of conscious effort to give associations that do not show the slightest resemblance to the accusing words (red, khaki; waste, stuff)

Presence of irrelevant replies coupled with extreme hesitation (spilt, sad)

Confession of inability to throw "off the scent" in giving no answer (leather, no reply)

Suspicious haste in making a reply of marked inconsequence (paper basket, yes, 1.4 seconds)

Stimulus Word	Reaction Words and Time in Seconds of Subject A		Reaction Words and Time in Seconds of Subject B	
	First Time	Second Time	First Time	Second Time
big	small 1.6	small 1.6	small 1.4	small 1.4
red	*green* 3.8	khaki 4.0	blue 1.2	blue 1.2
clock	tick 1.4	tick 1.4	time 1.2	time 1.2
run	jump 1.4	jump 1.4	walk 1.4	walk 1.4
brown	*bag* 1.4	leather 3.6	bag 1.4	bag 1.4
man	woman 1.4	woman 1.6	woman 1.2	woman 1.2
tick	tock 1.2	tock 1.2	tock 1.2	tock 1.2
leather	belt 3.2	no reply	bag 1.2	bag 1.2
bring	carry 2.2	carry 2.8	fetch 1.6	fetch 1.6
tree	field 1.6	field 1.8	branches 1.4	branches 1.4
paper basket	*down* 4.4	yes 1.4	paper 1.2	paper 1.2
table	chair 1.4	chair 1.4	chair 1.2	chair 1.2
house	room 2.0	room 1.4	street 1.4	street 1.4
waste	*rubbish* 3.0	stuff 4.8	paper 1.2	paper 1.2
sing	play 1.4	play 1.2	dance 1.4	dance 1.2
given	blue 1.2	blue 1.4	field 1.6	field 1.6
spilt	*dirty* 5.2	*sad* 4.2	milk 1.4	milk 1.6

This test is susceptible of endless modification and may prove highly revealing.

3. It is important to ascertain the subject's pedagogical standing. Mateer * makes the significant statement that: " The psychopathic case almost always succeeds in school, about 20% scoring below the grades they should cover. The test findings on school work are not sufficiently diagnostic to differentiate an individual unless the ratings are compared with mental age and

* *The Unstable Child*, Appleton, 1924.

other aspects of his ability, but there is a definite indication that, when mental age is fairly normal and school tests are low, the probabilities are three to one that the case is psychopathic rather than normal."

4. Healy stresses the diagnostic value of the subject's own story as detailed to the examiner. The account is sometimes startlingly revelatory, exposing to the view areas of behavior unexplored by any known test. He discloses a bewildering welter of myths, prejudices, motives, delusions, and infantile reasonings. He indulges in mock heroics, assumes a dramatic posture, or wallows abjectly in the mud of despondence. All these disclosures point to a definite answer.

5. Finally, the psychologist should give character tests and tests of emotional stability. The following tests may be found useful. They are reactive tests and to that extent do not establish permanency of traits, but they indicate prevailing moods.

Social relations — Ream
Emotional inhibitions — Crane
Emotional stability — Woodworth
Emotional affectivity — Pressey
Persistence — Chapman
Volitional perseveration — Downey
Suggestibility — Brown
Social-ethical vocabulary — Kohs
Content of moral concepts — Brotmarkle
Ascendance, submission — Allport

E. The School and the Neurotic

Having ascertained the extent of the disturbance, the school may set about to reëducate the child along social lines. The work, of course, is mostly extra-curricular and individual. Mass education is bound to fail. Each child must receive special treatment. If the trouble is physical, the family must be won over to a regimen aiming at the restoration of physical vigor and the rectification of such defects as poor teeth, diseased tonsils, nutritional derangements, poor eyesight, etc. Where the disorder is emotional the teacher should win the confidence of the pupil with a view to the gradual building up of self-confidence, perseverance, and the social amenities involved in the give and take of coöperative living. If the deviation is idiosyncratic, the aid of other children should be enlisted to discourage teasing and bullying. The school's contribution may be summarized in the phrase "sympathy without sentimentality." The more difficult cases need expert treatment of psychiatrists; there is a limit to the teacher's helpfulness. The education of neurotics cannot be undertaken as an incident of teaching. It has a technique of its own, and unfortunately its principles have not yet been formulated. There is still the gulf to bridge between institutional and tutorial methods of dealing with neurotics.*

* The "Exercises for Self-Testing and Investigation" which begin on page 305 are planned for use with this chapter also.

Bibliography

Army Mental Tests, Office of the Surgeon General, Washington, 1918

Ayres, L. P. — *Laggards in Our Schools*, Russell Sage Foundation, 1909

Baldwin, B. T. A. — "A Measuring Scale for Physical Growth and Physiological Age," *Fifteenth Year Book*, Part 1, National Society for the Study of Education, 1916

Ballard, E. F. — *An Epitome of Mental Disorders*, Blakiston's Son, Philadelphia, 1917

Barr, M. W. — "The How, the Why, and the Wherefore of the Training of Feeble-Minded Children," *Journal of Psycho-Asthenics*, September, 1899

Binet, A., and Simon, T. — *The Development of Intelligence in Children*, translated by E. Kite, The Training School, Vineland, N. J., 1916

—— *The Intelligence of the Feeble-Minded*, translated by E. Kite, The Training School, Vineland, N. J., 1916

Bolton, T. L. — "The Growth of Memory in Children," *American Journal of Psychology*, 1892

Bonaterre, A. — *Notice Historique sur le Sauvage de l'Aveyron*, Paris, 1799

Bronner, A. F. — *The Psychology of Special Abilities and Disabilities*, Bobbs-Merrill, 1917

California State Board of Charities and Corrections — *Surveys in Mental Deviation*, Sacramento, 1918

Case, I. — "The Correlation between Mental Defect and Anomalies of the Hard Palate," *American Journal of Insanity*, April, 1919

Cotton, H. A. — *The Defective, Delinquent and Insane*, Princeton University Press, 1921

Crafts, L. W. — "Bibliography of Feeble-Mindedness in Its Social Aspects," *Journal of Psycho-Asthenics*, Monograph Supplement, 1917

Doll, E. A. — *Anthropometry as an Aid to Mental Diagnosis*, The Training School, Vineland, N. J., 1916

Doll, E. A. — *Clinical Studies in Feeble-Mindedness*, R. Badger, Boston, 1917

Glueck, B. — "A Study of 608 Admissions to Sing Sing Prison," *Mental Hygiene*, January, 1918

Goddard, H. H. — "Height and Weight of Children in Relation to General Intelligence," *Pedagogical Seminary*, 1914

—— "The Menace of Mental Deficiency from the Standpoint of Heredity," *Boston Medical and Surgical Journal*, August 24, 1916

—— *Psychology of the Normal and Subnormal*, Dodd, Mead, 1919

—— "Two Thousand Normal Children Tested by the Binet Measuring Scale of Intelligence," *Pedagogical Seminary*, 1911

Holmes, A. — *The Conservation of the Child*, Lippincott, 1912

Hollingworth, H. L. — "Mental Ability and Chronological Age in Adults," *Psychology of Functional Neuroses*

Hollingworth, L. S. — *The Psychology of Subnormal Children*, Macmillan, 1924

Horn, J. L. — *The Education of Exceptional Children*, Century, 1924

Irwin, E. A. — *Truancy*, Public Education Association, New York, 1915

Itard, J. — *De l'Education d'un Homme Sauvage*, Paris, 1801

Kelly, R. L. — "Psychological Tests of Normal and Abnormal Children," *Psychological Review*, 1903

Kuhlmann, F. — *What Constitutes Feeble-Mindedness*, R. Badger, Boston, 1917

Mateer, F. — *The Unstable Child*, Appleton, 1924

McLearn, M. — "The Fact of Personality in the Development of a Typical Child," *Pedagogical Seminary*, March, 1913

Merrill, M. A. — "The Abilities of the Special Class Children in the Three R's," *Pedagogical Seminary*, March, 1918

Montague, H. — *Report of the Psychopathic Clinic of the Children's Court*, New York, 1918

Norsworthy, N. — "The Psychology of Mentally Deficient Children," *Archives of Psychology*, 1906

Ordahl, L. E. and G. — "Qualitative Difference between Levels of Interest in Feeble-Minded Children," *Journal of Psycho-Asthenics*, June, 1915

Peterson, A. M., and Doll, E. A. — *Sensory Discrimination in Normal and Feeble-Minded Children*, Training School Bulletin, November and December, 1914

Pintner, R. — *The Mental Survey*, Appleton, 1918

Pintner, R., and Paterson, D. G. — "Psychological Basis for Diagnosis of Feeble-Mindedness," *Journal of Criminal Law and Criminology*, May, 1916

Poull, L. — "The Mental Status of Truants" *Ungraded*, October, 1919

Schlapp, M. G., and Hollingworth, L. S. — "The Mentally Defective as Cases in the Courts of New York City," *Medical Record*, February, 1915

Sollier, P. — *Psychologie de l'Idiot et de l'Imbecile*, F. Alcan, Paris, 1901

Stenquist, J. L., Thorndike, E. L., and Trabue, M. R. — "The Intellectual Status of Children Who Are Public Charges," *Archives of Psychology*, 1915

Terman, L. M. — *The Rate of Growth and the Validity of the I.Q.*, Warwick and York, 1917

—— *Research in Mental Deviation among Children*, Stanford University Bulletin, No. 2, 1915

—— *Sex Differences*, Warwick and York, 1911

Three Problem Children, Joint Committee of Preventing Delinquency, 50 E. 42d St., New York City

Tredgold, A. F. — *Mental Deficiency*, William Wood & Co., 1915

Wallin, J. E. W. — *The Education of Handicapped Children*, Houghton Mifflin, 1924

—— *Problems of Subnormality*, World Book, 1917

Watts, F. — *Abnormal Psychology and Education*, Appleton, 1924

White, W. A. — *Principles of Mental Hygiene*, Macmillan, 1917

CHAPTER X

THE SUPERIOR CHILD

A. Attitude toward the Superior Child

(1) *Attitude in the Past*

Before the advent of the mental tester, the attitude toward the superior child oscillated between scorn and indifference. Genius was popularly accounted as potential insanity; and intrenched mediocrity frowned upon superiority. Many psychiatrists viewed precocity with an interrogation as to its mental stability. The school was partial to the superior child for administrative reasons. His presence raised the scholastic tone of a class, and his phenomenal progress acted as a goad to the slower pupils.

However, not every gifted child took kindly to such exploitation of his powers. The exigencies of mass teaching slowed down the teaching tempo to a pace commensurate with the comprehension of the less gifted. He became listless and mischievous, a difficult case. The distinction of being a paragon of learning was not conducive to popularity, especially when the teacher's unstinting praise was followed by satiric references to the obtuseness of the less favored ones. Ostracism or persecution was often the lot of the supe-

rior child, with consequences disastrous to his emotional integrity. When the gifted child became a disciplinary problem, his condition was deplorable in the extreme. His very versatility and originality were suspected. He encountered hostility at every turn, and the bushel of discouragement was thrust prematurely upon the light of his genius. Many lights were thus summarily snuffed out. More hardy souls removed this handicap by setting fire to the bushel. History abounds in examples of great men who were forced to wage relentless war against pedagogical coddling no less than against pedagogical starvation. Their biographies are a series of jail deliveries from the preceptorial prisons into which bungling schoolmasters consigned them.

In their treatment of the gifted child the schools have erred as grievously as in the mismanagement of the subnormal.

(2) *Altered Attitude in the Present*

The present attitude may be defined as unsentimental training of the gifted, untrammeled by the feeling of reverence or pity — awe for their precocity or sympathy for their fate. Educators are disposed to disallow the claims of some that precocity is a presumption of insanity or of ill health. It is freely admitted that the condition of mental and physical health varies among the superior. An unselected group of gifted children is bound to include individuals predisposed to disease, but the middle fifty per cent are free from all

taints and may be educated without special regard to stigmata supposedly latent in their constitution.

B. Characteristics of the Superior Child

The outstanding characteristic of the superior child is an abundant but restless energy flowing in full current in all channels of personality. He is not merely more intellectual than the normal. He excels in all four dimensions of persoanlity: temperament, intelligence, morality, energy reservoirs. Unlike the defective, whose personality is truncated, or the normal, with a more dynamic personality exhibiting tendencies to compensate inferiority in any personality traits by greater competency in others, the superior child faces the world in the fullness of all powers. To employ a trite figure of speech, the defective has few psychic assets, with a penchant for hoarding or plunging; the normal child is blessed with a fair amount of capital, which he invests conservatively; the superior child has unlimited capital, in the form of liquid assets, which he is prepared to invest in more speculative enterprises. True, his judgment is not always safe, but he rarely gambles without resort to some system and his indomitable courage in the face of failure hastens his rehabilitation. Even bankruptcy does not extinguish his enthusiasm or dampen his ardor.

When the superior child is segregated into homogeneous groups, so that he is not floundering out of his depth, dallying in the shallows, or sulking on the bank,

he is tractable, social, honorable, and energetic. He gets along capitally with his fellows and teachers, learns quickly and recalls as rapidly, displays intellectual compromises, and cleaves to a line of conduct with a force of energy that does not abate with failure. Contrary to the popular notion, he is not complex-ridden, thanks to his extraordinary flexibility. In fact, his ability to meet new problems with open mind and unprejudiced attitude, without trepidation or overweening confidence, is as characteristic of his personality as his abounding vitality.

(1) *Studies of Superior Children*

In 1922 Yates made a study of two high school groups, 25 average pupils and 25 superior pupils. They were grouped on the basis of intelligence tests. The superior group excelled in academic marks and ranked higher in the estimation of teachers. Other differentiating qualities were these: they (1) walked and talked earlier; (2) matured earlier; (3) learned to read and write earlier; (4) had a greater liking for the school; and (5) were voracious readers. Yates's study discloses the fact that superior children are recognizable by other signs than the I.Q.: the possession of energy sources adequate to meet the demands of school life, with ample reserves for extra-school activities; these springs of energy well up in early childhood and never go dry.

In 1919 Terman reported a study of 59 children with

an I.Q. above 136. They ranged between the ages of three and fourteen. His findings are interesting: they are far above the average in "moral and personal traits"; their social adaptations are adequate.

The Germans have displayed a keen interest in the superior child. Stern has studied the problem intensively. He recommends that the selection of superior children should be made on the basis of (1) the teacher's judgment, (2) scholastic record, (3) the psychograph of the pupil's personality traits, (4) achievement tests, and (5) an intelligence test. The last item needs explanation. The Germans have pinned their faith in tests of discrete psychic powers rather than in tests of general intelligence. While they do not espouse the obsolete faculty psychology, they like to think of intelligence as constituted of observation, memory, reasoning, and imagination. Their tests, therefore, are designed to evaluate each one of these aspects of intelligence, expressing the result in terms of the total score attained in each.

Coy notes the improvement in scholarship, but enthusiasm is reserved for the recital of new attitudes displayed by the pupils. Here were evident a spirit of coöperative endeavor, a desire to shoulder responsibility, an awakening of powers dormant through lack of competition, a spirit of emulation without malice, and a realization of the painful fact that the feeling of superiority that had been fostered by a too facile victory over his less gifted mates had been built upon an inse-

cure foundation. With realization came short-lived humiliation, which was quickly succeeded by the resolve to fortify himself against failure by close study and sedulous practice. In other words, superiority manifested itself in the realignment of all the forces of personality and not in a mere changing of the mental pace, for intelligence by itself does not produce superiority. It works only in conjunction with a harmonious development of other traits of personality.

The most thorough investigation was reported by Terman in his *Genetic Studies of Genius*, Vol. 1. Between 1921 and 1925, he and his assistants made a statistical study of over 1000 gifted children. The final study embraced only 643 children, who were to be subjected to a comprehensive system of follow-up over a period of years. He finds that superior children follow in their distribution the curve of probability, varying from brightness to genius and grouping themselves in insensible gradations around the median; their achievements are not limited to any department of school life; they are equally successful in tests of character; they show less predisposition to " nervousness " than normal children; and they shoulder responsibility willingly and cheerfully. They not only study concentratedly, but they play hard, ride many hobbies, read more and better books, excel in athletics, and are animated by the highest sentiments of honor and integrity.

Catherine M. Cox has made a historical study of 300

geniuses with a view to reconstructing their boyhood. It is entitled *Genetic Studies of Genius*, Vol. 2. She employs a psychographic evaluation of desirable traits, which are rated on a scale of 7. There are twelve traits:

1. Persistence of motive
2. Intellectual activity
3. Positive self-traits
4. Strength or force of character
5. Intellectual traits
6. Social traits
7. Intellectual-social traits
8. Balance
9. Emotional-social traits
10. Physical activity
11. Emotional traits
12. Negative self-traits

Many of the traits are self-explanatory. Others need elucidation. In "positive self-traits" are included such motives as the desire to excel, self-confidence, and desire to dominate. "Intellectual traits" are a composite of originality, common sense, profoundness, memory, quickness, and observation. "Social traits" and "intellectual-social traits" are refinements of classification that lead to overlapping or quibbling. So do "emotional-social traits" and "intellectual-social traits." It is quite possible to dissect the personality into infinitely small unit-traits, but it is ques-

tionable whether the human personality can be thus anatomized and still conserve its unitary aspect.

(2) *Methods of Selecting and Grouping Superior Children*

Miss Lulu M. Stedman organized * in 1918 an opportunity room for gifted children in the training school of the University of California, Southern Branch. In this room were placed a number of gifted children who had been tutored at home or at school. In 1922 it became necessary to separate the seniors from the juniors, and another opportunity room was organized. Entrants were given both intelligence and achievement tests. Among the intelligence tests given were the Stanford-Binet, the Army Alpha, and the Terman Group Test of Mental Ability. Standardized educational tests supplemented the findings of the mental tests. Great dependence was placed upon the Stanford Achievement Test. The miscellaneous tests included:

Monroe Silent Reading Test III, Form 1
Thorndike-McCall Reading Scale, Form 3
Woody-McCall Mixed Fundamentals, Form II
Trabue Language Test, Form I
Charters Diagnostic Language Test, Form I
Modified Ayres Spelling Test

For five cases, which Miss Stedman describes in detail, she obtained other information derived from "baby-books, statistics compiled by parents, social histories, and the writer's observation."

* *Education of Gifted Children*, World Book Co., 1924.

The membership of the opportunity rooms comprised 16 cases. While a few children fell below the minimum I.Q. (125 for 1918 and 140 for 1922), the pupils exceeded the norms in all school subjects. "In both reading tests and in the Trabue Language Scale B, the mean of the group exceeds that of high school students about to graduate. In all other subjects the mean is higher than the norms for the eighth and ninth grades."

Her findings accord with those of other investigators, and her plea for the continuance and extension of opportunity rooms is based upon the fact that (1) progress is consistent with ability, (2) discipline becomes a builder of character, (3) the asocial become socialized, (4) exceptional talent is afforded scope for development, and (5) the smallness of the group conserves the intellectual benefits of competition without its accompanying antisocial attitudes.

Another method designed to meet the exigencies of large systems is the one used in the Isaac Remsen Junior High School of Brooklyn, N. Y. This school receives for rapid advancement classes pupils who are differentiated in the 7A grade. They are henceforth known as RA, RB, RC, and RD classes, which cover the work of the 7A, 7B, 8A, 8B, 9A, and 9B (9A and 9B, first year of high school) in two years. The RA pupils are recruited from among the best 6B pupils of neighboring schools and are chosen by their respective principals. They are received into the junior high school on the recommendation of their principal. Some

of these entrants are superior children; some are bright children; while not a few are on the borderline between normal and bright. The latter are given a fair trial and are either assigned to the lowest RA or are eliminated at promotion time; some go to 7B, others to 8A.

Upon their assignment to a class, all RA pupils are placed under the supervision of an adjustment teacher, whose duty it is to test the pupils at set intervals (usually at the beginning of every term) and place them to their best advantage. This teacher acts as liaison officer between teacher and pupil and teacher and principal. His activities carry him far afield, making exploratory studies of courses of study, extra-school interests of pupils, and vocational plans of graduates. The greatest task set before him is that of adjusting the scholastic load to the shoulders of the pupils and of energizing the indolent or the inactive.

The process of adjustment is set in motion immediately following the assignment of new entrants to classes. The basis of their classification is tentative and at times arbitrary. The presumption is that all are competent to do the advanced work of an enriched curriculum. This period of trial, and in a sense incubation, lasts ten weeks. In the meantime data are being gathered from the following sources:

1. Combined scores are obtained in group intelligence tests. The Terman Group Test of Mental Ability has proved itself highly discriminating, though the National

Intelligence Test and the Otis Intermediate Self-Administering Test have been used as correctives.

2. Standardized achievement tests and traditional tests (tests in class work) have been checked against each other to establish the causes for unsatisfactory results. A pupil whose scores in achievement tests equal or exceed the norm for his grade may fail signally in class work because (1) he is physically indisposed, (2) he is not applying himself with sufficient diligence, (3) he is irregular in attendance, (4) he is employed in a gainful occupation after school and is forced to neglect his home work, (5) he is not getting along with the teachers, (6) the teacher has not adapted her instructional and disciplinary methods to superior children, or (7) he cannot stand the pace set by the rest of the class.

The pupil who fails for any reason or reasons here enumerated is subjected to an exhaustive investigation involving the use of tests of character and emotional stability, a physical examination wherever parental coöperation can be secured, a detailed report by the teacher, an interview with parents, an interview with the pupil, a daily examination of class work, and a conference in which all teachers to whom the pupil recites participate.

3. The pooled judgment of all teachers is another source of information. This judgment is based upon daily observation of other facets of personality in addition to intelligence. The teacher's report includes (1) class standing, (2) idiosyncracies that hinder prog-

ress, (3) extra-curricular activities, (4) mental and emotional attitudes, (5) intimations of emotional instability, (6) the presence of sensory defects, (7) the degree of gregariousness displayed and whether inability to get along with his fellows is due to introversion or to temperament, and (8) the ambitions and aspirations cherished in the recesses of his mind. The reports of the various teachers are examined for signs of bias or evident dislike, and allowance is made for the vagaries of personal opinions.

4. The report cards are studied for the entire school career to note attendance, progress, and ratings. A rating of B+ in effort, proficiency, and conduct for the two years preceding promotion to the 7A is deemed a prerequisite, though a lower rating does not, in itself, operate as a bar to admission to the RA organization.

The adjusting teacher is thus placed in possession of information sufficiently comprehensive to warrant a tentative reclassification into superior, bright, and unadjusted. The superior are henceforth known as RA, the bright as RA2, and the unadjusted as RA3. However, the classification is not final. Not a few of those adjudged superior or bright become in time decidedly maladjusted and gravitate to the RA3, and members of the RA3 may develop unsuspected powers or may make up in assiduity what they lack in capacity. Such conditions having been established, the pupil is transferred to the appropriate group. The process of adjustment continues to the point of satisfactory place-

ment, within the RA organization or outside of it. It would seem that the bright and gifted, by reason of the plasticity of their personality traits and the inflexibility of the school organization, are more liable to become maladjusted than the children of moderate endowments.

C. Education of the Superior Child

Having discovered the manifold aptitudes of superior children, educators set about enlarging their educational opportunities. While in theory the projectors of new curricula widened the boundaries of both cultural and factual subjects, in practice courses were inflated to encyclopædic scope and the "enrichment of the curriculum" led to the resurrection of topics and processes long discarded as superfluous, irrelevant, or technological. Education degenerated into training, and instruction became vocationally purposive or was directed toward preparation for high school or college. So awed were teachers by the spectacle of genius in the raw state that they were tormented by the fear lest by their ineptness they might either destroy the germ or retard its germination. They therefore bent all efforts to rush the gifted child through the grades in order that the process of maturation might proceed under more auspicious surroundings. It was mistakenly argued that the elementary school had nothing of value for the superior child and that his attendance was a mere formality; he did not tarry beyond the time it took

to comply with the requirements for entrance to high school.

But the speeding-up process neither fostered genius nor enhanced scholarship. For one thing genius, like normality, follows a definite rhythm of development and cannot be forced with impunity. For another this frantic amassing of facts but fills one compartment of the personality and disturbs the harmony that obtains between the minor forces of personality. You cannot persist in the practice of filling the mind as you would fill a library without imparting to it the oppressive atmosphere of yellowing leaves and crumbling buildings; and in an atmosphere of stifling scholarship the social graces cannot thrive and emotions become too intellectualized.

This deflection of the turbulent current of personality into a narrow channel of intellectualism has resulted neither in deepening the bed nor in widening the channel. The new pedagogy is combating this tendency to sterility. Its principles are based on the psychology of the superior child. It is effectuating modifications of curriculum and methods and is evolving a technique of its own.

(1) *Modifications of the Curriculum*

The curriculum should be modified in the following directions.

1. Depth. This does not mean enriching the curriculum by new accretions of facts. It has regard to

the presentation of a topic that would lead to a discovery of its uttermost ramifications consonant with the mental status of the learner; fixing its position in its immediate hierarchy and relating it logically to allied topics and to the subject as a whole; exploring its implications for scholarship and civilization and discouraging a too facile abridging of the gap between demonstrable proof and plausible inference.

2. Thoroughness. Thoroughness may be achieved in a very narrow field, but the thoroughness here commended refers to a mastery of a subject in both letter and spirit. It was the straining after such completeness that led Upton Sinclair to apply to the college authorities for a two months' leave of absence in order to prosecute his studies with more vigor and continuity. It was near the end of the senior year that Sinclair decided to break the shackles of lockstep education. He spent the two months " sitting on a bed in an eight by ten hall bedroom, in a lodging-house, reading Shelley's poetry and Emerson's essays and the prose of Ruskin and Carlyle."

3. Æsthetic Training. The appreciation of art is not of spontaneous growth; nor is it solely an emotional capacity. A great deal of art is intellectual. It requires for its fullest enjoyment training in such technical matters as color, design, craftsmanship, and practice in distinguishing nuances of sound or color. To neglect art is to deprive the gifted child of one more adequate release for his abundant energies and to condemn him

to divert them into streams too shallow for adequate utilization. Besides, art is recreational in its essence, and leisure plays a great rôle in the life of the superior child, for his capacity for work is enormous and he easily confounds work with play to the detriment of his physical and emotional well-being.

(2) *Modification of Methods*

The technique of teaching superior children partakes of the characteristics of both mass and tutorial instruction. There are many activities in which the group may be taught as a whole; but the versatility of the group is so varied that it is difficult to find more than a few whose interests run along rigidly parallel lines. There is a marked divergence from pupil to pupil. Individual guidance becomes the prevailing practice, and personal contact of the most intimate kind must be established between pupil and teacher. The technique of individual instruction then calls for the marginal control of masses, with the individual as the focus of instruction — for an acceptable compromise between the rigidity of class discipline and the freedom of the tutorial organization. In the choice of methods, emphasis should be placed on (1) rationalization, (2) mastery of method rather than of rules, (3) originality, and (4) individual systems of memorization.

1. Rationalization. The tendency to defer rationalization of involved processes, in the belief that with increased maturity comprehension will prove more lasting,

cannot be justified in the training of superior children. Sinclair's resentment against the dulling effects of mechanization of method is shared by many. He inveighs against the practice of stifling exuberant inquiry in a straitjacket of rules. "If you wanted to divide one fraction by another, you turned the second fraction upside down; it seemed an odd procedure, but if you asked the reason for it, the teacher would be apt to answer in a way that caused the other boys to laugh at you — something which is very painful." The schoolboy of to-day rarely experiences such humiliation, but the logic underlying mere processes is not always elucidated unless the explanation may be made with axiomatic clearness. Therefore, teachers who mistake plausibility for logic and acquiescence for understanding are not favorably impressed with the results of rationalization. But the gifted child is rarely satisfied with less than the logical explanation and is apt to temper his respect for the teacher with mental reservation as to his ability and willingness to set at rest the fretting doubts. With few exceptions, the reasons for doing things may be revealed to the gifted child if the teacher will call a halt to ostentatious inquiries or doubts raised for the mere voicing of them. That logic has a way of disclosing increasingly abstruse reasons if followed to great lengths is no reason for reserving it. That logic has a way of betraying the unwary into fallacies or even absurdities does not warrant the exercise of excessive circumspection. It is precisely against this fatal lapse

into irrationality that the gifted must guard himself and the mass of the people.

2. Mastery of Method. The superior child must be put into possession of the more elementary principles of the psychology of learning, uncomplicated by technical terms or metaphysical implications. Not only should processes be rationalized, but the teacher must explain the reasons for pursuing definite systems of presentation, memorization, and drill. Take the subject of spelling. The superior child should know how multiple sense appeal aids memorization, and how adequate pre-testing and post-testing save both time and energy. It would not be amiss to explain in detail how to attain mastery over " demonic " words, such as " Wednesday " and " always," and such type or key words as " stopped," "getting," "writing," and "opened." Of course mastery of method, like rationalization, may readily become an end in itself to the neglect of factual material and the pernicious inflation of the pupil's ego. A situation may arise in which a pupil accumulates a bag full of tricks for learning speedily without putting any one to profitable and immediate use; not unlike the ineptness of the author who publishes a book on " A Hundred Ways of Getting Rich " and starves in an attic. But between the two evils — mastery of facts without mastery of method and mastery of method with imperfect mastery of facts — it is not difficult to make a choice. The practical teacher can readily make an equitable adjustment.

3. Originality. As far as it is consistent with preparation for a definite goal — regents or college entrance — superior children should be afforded an opportunity to exercise their creative faculties. The teacher must be heedful of the logical and practical limitations of originality. The world can well spare some of its original thinkers who spin shimmering webs, whose warp is imagination, and whose woof is phantasy. But of thinkers and artists who soar into the realm of the imaginative and poetic without losing sight of the fact that the rarefied atmosphere they breathe must be made fit to sustain the less adaptable members of the race, the world has a dearth and the supply of such men must be replenished with drafts from the superior children. Not that originality may be made the subject of instruction. Originality, however, may be made amenable to pedagogic rules by conforming to its laws of development and by applying the criteria of common sense, practicability, and utility.

4. Inventiveness in Devising Ways of Memorizing. The capacity to remember facts and processes is highly developed in the superior child, and this virtue may be credited to the ability to multiply associative threads on which facts are strung. Indeed the gifted children surpass the less favored individuals in the ease with which they can cause two associative strands to grow where only one grew before. This enhanced fertility is due to an acute awareness of the nexus that binds objects, no matter how dissimilar in their superficial

aspects. When logical cement is wanting, recourse is had to factitious relations expressed in the form of mnemonics. It may prove a revelation to many teachers to learn what resourcefulness bright children display in constructing devices for effecting greater cohesiveness where repulsion supposedly obtains. Like all activities of the superior intellect, the need of curbing the exuberant fancy is imperative, for the mnemonics may easily degenerate into a meretricious artificiality burdening the memory with associative bonds that do not grow naturally from the material to be memorized. However, individual variability should be considered. Some minds develop unique associative tentacles that enable the possessor to grip and hold facts that would make no impression on others. The condemnation of the mnemonic system should not be made on purely a priori grounds; the test of efficacy should determine its validity.

Miss Stedman summarizes the modifications essential for the training of gifted children as follows:

1. A curtailment of the time allotment of the various subjects to from one-half to two-thirds the time normally assigned
2. Reduction of drill, explanation, and development
3. Definite instruction in how to study independently
4. Opportunities for all forms of creative work — the writing and presentation of plays, pageants, and pantomimes

5. Definite training in leadership
6. Special courses in music, art, and dramatics
7. Manual work possessing educational value
8. Opportunities for recreational activities
9. Development in the child of a scientific attitude toward his own progress and of a scientific method of attacking problems *

BIBLIOGRAPHY

Almack, J. C. and J. L. — "Administrative Problems Connected with Gifted Children," *Educational Administration and Supervision*, Vol. 8, No. 3, pp. 129–136

Cleveland, E. — "Some Further Studies of Gifted Children," *Journal of Educational Research*, Vol. 4, No. 3, pp. 195–199

Counts, G. S. — "Social Purpose of the Education of the Gifted Child," *Educational Review*, Vol. 64, p. 233

Courtney, W. L. — *Life of John Stuart Mill*, London, 1889

Coy, G. L. — "The Mentality of a Gifted Child," *Journal of Applied Psychology*, Vol. 2, pp. 297–307

Downes, F. E. — "Seven Years with Unusually Gifted Pupils," *Psychological Clinic*, Vol. 6, p. 13

Garrison, C. G. — "The Psychology of a Prodigious Child," *Journal of Applied Psychology*, Vol. 1, pp. 101–110

Gillingham, A. — "Superior Children — Their School Progress," *Journal of Applied Psychology*, Vol. 9, No. 6, pp. 327–347

Henry, T. S. — "Class-Room Problems in the Education of Gifted Children," *Nineteenth Year Book*, Bloomington, Illinois, 1920

Hirt, Z. I. — *A Gifted Child*, Training School Bulletin, Vol. 19, p. 49

Hollingsworth, L. S. — *Gifted Children*, Macmillan, 1926

Irwin, E. A., and Marks, L. A. — *Fitting the School to the Child*, Macmillan, 1924

Peters, C. C. — *Foundations of Educational Sociology*, Macmillan, 1924

* The "Exercises for Self-Testing and Investigation" which begin on page 305 are planned for use with this chapter also.

Race, H. V. — "A Study of a Class of Children of Superior Intelligence," *Journal of Educational Psychology*, Vol. 9, No. 2, pp. 91–98

Root, W. T. — *A Socio-Psychological Study of Fifty-Three Super-Normal Children*, Psychological Monograph, No. 133

Stedman, L. M. — *Education of Gifted Children*, World Book, 1924

Stern, W. — "The Supernormal Child," *Journal of Educational Psychology*, Vol. 2, pp. 143–148, 181–190

Terman, L. M. — "The Intelligence Quotient of Francis Galton in Childhood," *American Journal of Psychology*, Vol. 28, pp. 209–215

—— "A New Approach to the Study of Genius," *Psychological Review*, Vol. 29, No. 4, pp. 310–318

Terman, L. M., and Fenton, J. C. — "Preliminary Report of a Gifted Juvenile Author," *Journal of Applied Psychology*, Vol. 5, No. 2, pp. 163–178

Town, C. H. — "Superior Children in Our Schools," *Educational Review*, Vol. 65, p. 17

Whipple, G. M. — *Classes for Gifted Children*, Public School Publishing Co., Bloomington, Illinois, 1919

Witmer, L. — "The Training of Very Bright Children," *Psychological Clinic*, Vol. 13, p. 88

Witte, K. — *The Education of Karl Witte*, Translated by L. Wiener, Crowell, N. Y., 1914

Yates, D. H. — *A Study of Some High School Seniors of Superior Intelligence*, Journal of Educational Research Monograph, No. 2

Zirkle, H. W. — "Taking Care of the Gifted Child," *First Year Book*, Department of School Principals, National Education Association

FINAL TEST

In each of the following questions three references are given. Select the one that is correct or is generally accepted to be best. It is preferable to make no selection in those questions where you lack knowledge completely as incorrect answers will be penalized.

1. Intelligence:
   (*a*) Stated qualitatively
   (*b*) Stated quantitatively
   (*c*) Stated descriptively
2. Intelligence:
   (*a*) A measure of the pupil's capacity for the work of the school
   (*b*) Native learning ability
   (*c*) Capacity to think abstractly
3. Intelligence:
   (*a*) Independent of learning
   (*b*) Independent of degree of retentiveness
   (*c*) Independent of special abilities and talents
4. A feeble-minded child
   (*a*) may be prepared for college.
   (*b*) may be prepared for high school.
   (*c*) may hope to reach the fourth grade.
5. General common factor in intelligence:
   (*a*) Burt
   (*b*) Stern
   (*c*) Goddard
6. Vocabulary tests have been used as
   (*a*) composition tests.
   (*b*) tests of trustworthiness.
   (*c*) tests of intelligence.

7. Test not normally given in school:
   (*a*) "directions test"
   (*b*) "vocabulary test"
   (*c*) "make one sentence which includes, as an integration, three specified words."
8. Most tests of intelligence yield
   (*a*) mental ages.
   (*b*) coefficients of brightness.
   (*c*) point scores.
9. Directions for giving most tests are found in
   (*a*) the test booklets.
   (*b*) special "directions" blanks.
   (*c*) manuals.
10. The I.Q. forms a basis for predicting
    (*a*) school success.
    (*b*) social effectiveness.
    (*c*) level of intelligence that the subject will probably reach in the future.
11. Group tests for college students:
    (*a*) Binet
    (*b*) Woody
    (*c*) Thorndike
12. Bobertag
    (*a*) devised tests in reading.
    (*b*) prosecuted experiments in memory.
    (*c*) published a German revision of the Binet Scale.
13. Stanford Revision:
    (*a*) Goddard
    (*b*) Thurstone
    (*c*) Terman
14. A distinctive addition to the original Binet Scale:
    (*a*) "Counting backwards"
    (*b*) "Pointing to parts of body"
    (*c*) "Baseball field"

15. "What's your last name" (Stanford Revision):
    (*a*) Ninth year
    (*b*) Second year
    (*c*) Third year

16. I.Q. 90 (Binet):
    (*a*) Bright
    (*b*) Feeble-minded
    (*c*) Normal

17. "Nearly all who are below 70 or 75 I.Q. are feeble-minded":
    (*a*) Thorndike
    (*b*) McCall
    (*c*) Terman

18. The feeble-minded should be tested by
    (*a*) teachers.
    (*b*) psychiatrists.
    (*c*) psychologists.

19. The Stanford-Binet Scale is less effective in measuring
    (*a*) normal children.
    (*b*) young children.
    (*c*) superior older children.

20. A true measure of a child's intelligence is furnished by
    (*a*) one testing.
    (*b*) two testings.
    (*c*) one testing supplemented by the teacher's judgment.

21. The I.Q. is
    (*a*) absolutely constant.
    (*b*) relatively constant.
    (*c*) approximately constant.

22. Addition of tests for children under three years of age:
    (*a*) Burt Revision
    (*b*) Herring Revision
    (*c*) Kuhlmann Revision

23. Has made distinctive contributions to the study of character testing:
    (*a*) Thorndike
    (*b*) Jung
    (*c*) Fernald
24. Has made an exhaustive study of the delinquent:
    (*a*) Otis
    (*b*) Wallin
    (*c*) Monroe
25. The Stanford Revision has been severely criticized for making undue use of
    (*a*) numbers.
    (*b*) pictures.
    (*c*) language.
26. "Tapping Test":
    (*a*) A linguistic test
    (*b*) A construction test
    (*c*) A performance test
27. Pintner-Paterson Performance Scale depends on
    (*a*) no language.
    (*b*) pantomime.
    (*c*) a minimum of language.
28. A pioneer in the construction of group tests of intelligence:
    (*a*) Baker
    (*b*) Van Wagenen
    (*c*) Otis
29. A general power test:
    (*a*) Time limit
    (*b*) No time limit
    (*c*) Time left to discretion of the examiner
30. "Draw a line under the one word that makes the sentence true: Coffee is a kind of . . . bark . . . berry . . . leaf . . . root." This is a
    (*a*) directions test.
    (*b*) classification test.
    (*c*) information test.

31. Nonverbal intelligence examination:
    (*a*) Otis Self-Administering Test
    (*b*) Miller Mental Ability Test
    (*c*) Dearborn Group Test of Intelligence, Series I
32. Diagnosis in school achievement is
    (*a*) a testing process.
    (*b*) a process of psychological analysis.
    (*c*) an observational process.
33. Scales are mostly employed in testing
    (*a*) arithmetic.
    (*b*) handwriting.
    (*c*) oral reading.
34. The chief cause of poor work in arithmetic is
    (*a*) ignorance of fundamentals.
    (*b*) poor teaching.
    (*c*) poor mentality.
35. The best method of improving spelling efficiency is by
    (*a*) recourse to the Courtis Practice Tests in Spelling.
    (*b*) instilling a spelling "conscience."
    (*c*) frequent testing.
36. Made contributions toward the statistization of educational measurements:
    (*a*) Woodworth
    (*b*) Kohs
    (*c*) McCall
37. A spelling list may be constituted into a standardized measure of spelling ability provided
    (*a*) the choice is based on a study of many spelling books.
    (*b*) the words are selected by good teachers of spelling.
    (*c*) the difficulty value of each word is ascertained.
38. "Composite score":
    (*a*) Average score of a number of tests
    (*b*) Median score of a number of tests
    (*c*) Average of the score of several tests expressed in terms of a common unit

39. Difficulty value of a test item is usually expressed in terms of
    (*a*) percentage of successes made by a sampling of children.
    (*b*) P.E. value.
    (*c*) S.D. value.
40. "Normal distribution":
    (*a*) Skew-shaped
    (*b*) Oblong-shaped
    (*c*) Bell-shaped
41. Probable error (P.E.):
    (*a*) Measure of correlation
    (*b*) Device used to describe the variability or spread of the normal distribution about its average
    (*c*) Measure of skewness
42. Validity:
    (*a*) Degree of correlation with a standardized test
    (*b*) Correspondence of scores obtained in giving a test for the first time with scores obtained through its repetition
    (*c*) Truthfulness with which a test fulfills its purpose
43. The most important contributant to good scholarship is
    (*a*) personality of the teacher.
    (*b*) good methods of teaching.
    (*c*) good methods of studying.
44. Heredity and environment:
    (*a*) Contribute equally toward the formation of good character.
    (*b*) Heredity conditions character building.
    (*c*) Environment is the preponderating factor in character building.
45. That standardized test is to be given preference which has
    (*a*) many items in each subtest.
    (*b*) alternative forms.
    (*c*) high correlation with the Stanford-Binet Test.
46. The relative scholastic proficiency of a group is best indicated through
    (*a*) median.
    (*b*) percentile ratings.
    (*c*) interquartile measure.

47. The difference in general intelligence obtaining between groups of school children should be interpreted in terms of
    (*a*) school administration.
    (*b*) school management.
    (*c*) the psychology of learning.

48. Buckingham in *Research for Teachers* says that "the measurement of intelligence is indeed fundamental for school purposes. . . . Every child is born equal in the right to become all that his nature permits him to become. With reference to the school this imposes the obligation to study each child. First and foremost, it means the study of his intellectual endowment." This means that Buckingham considers non-intellectual traits
    (*a*) secondary.
    (*b*) inextricably bound up with intellectual traits.
    (*c*) derivatives from intellectual traits.

49. The tests and scales for measuring intelligence are, to a very large extent, based on
    (*a*) the original Binet tests.
    (*b*) the Army tests.
    (*c*) the Bourdon Completion Test.

50. The "precocious" child is in reality
    (*a*) a superior child.
    (*b*) a child "prematurely done."
    (*c*) one who is maturing at an abnormally rapid rate.

51. The most serious lack to-day in the field of intelligence tests is tests for
    (*a*) subnormal children.
    (*b*) normal children.
    (*c*) gifted children.

52. The Stanford-Binet is inadequate for measuring gifted children because
    (*a*) in the upper age levels of the scale the "steps" are too large (four, five, and six months of mental age).

(*b*) the tests are too easy for them.
(*c*) there is nothing in the tests to stimulate and hold the interest of gifted children.

53. Very superior children tend to
(*a*) progress beyond maturity.
(*b*) remain stationary.
(*c*) regress toward mediocrity as they reach maturity.

54. "A fundamental stimulus" is an element of the concept
(*a*) configuration.
(*b*) James-Lange theory of emotions.
(*c*) conditioned reflex.

55. "The Spearman foot-rule":
(*a*) A measure of intelligence level
(*b*) A measure of correlation
(*c*) An arithmetical formula in mensuration

56. Arithmetical prodigies are
(*a*) generally of low intelligence.
(*b*) usually highly gifted.
(*c*) found among both the bright and the dull.

57. Mechanical talent seems to be
(*a*) a concomitant of dullness.
(*b*) a specialized ability independent of the quality of intelligence.
(*c*) positively correlated, in many investigations, with both dullness and brightness.

58. The most comprehensive procedure for discovering gifted children is to be found in
(*a*) Berlin.
(*b*) New York City.
(*c*) London.

59. "Configurative learning":
(*a*) Trial and error
(*b*) S–R
(*c*) Gestalt

60. A good test of morality provides
    (*a*) an understandable situation.
    (*b*) a natural situation provocative of a natural response.
    (*c*) a situation whose ethical implications are recognized as coming within the moral code of the child's social set.

61. Vocational aptitudes are best discovered through
    (*a*) intelligence and achievement tests.
    (*b*) first-hand contacts with vocational processes under "shop" conditions conducted under school guidance.
    (*c*) self-discovery in the world of industry.

62. Emotions enter decisively into that phase of the teaching process known as
    (*a*) generalization.
    (*b*) abstraction.
    (*c*) motivation.

63. Investigators of gifted children usually present their findings in the form of
    (*a*) statistical studies of intelligence and achievement scores.
    (*b*) "cases."
    (*c*) graphs.

64. Psychopathic children are best educated in
    (*a*) small groups.
    (*b*) ungraded classes.
    (*c*) traditionally organized classes in charge of an "understanding" teacher.

65. Psychopathy is conditioned by
    (*a*) temperamental bias.
    (*b*) dementia præcox.
    (*c*) capricious up-bringing.

66. Dementia præcox affects
    (*a*) "shut-in" children.
    (*b*) foreign-born children.
    (*c*) anæmic adults.

67. The family history is of substantial aid in the diagnosis of
    (*a*) malnutrition.
    (*b*) psychopathy.
    (*c*) inattentiveness.

68. A pupil's achievement in silent reading (as tested by achievement tests) is
    (*a*) one-dimensional (speed).
    (*b*) three-dimensional (speed, comprehension-prose, comprehension-poetry).
    (*c*) two-dimensional (speed, comprehension).

69. Arithmetical abilities are
    (*a*) general.
    (*b*) hierarchized.
    (*c*) specific.

70. In arithmetic there are as many different abilities as there are
    (*a*) children.
    (*b*) types of examples.
    (*c*) bonds.

71. The highest positive correlation subsists between
    (*a*) addition and intelligence.
    (*b*) counting and intelligence.
    (*c*) division and intelligence.

72. Deviser of tests in arithmetical reasoning:
    (*a*) Thorndike
    (*b*) Buckingham
    (*c*) Ayres

73. Köhler:
    (*a*) Reading Scale
    (*b*) "Configurative learning"
    (*c*) Handwriting Scale

74. Combined intelligence-achievement test:
    (*a*) Otis Classification
    (*b*) Stanford Achievement
    (*c*) Myers Mental Measure

75. The most important element in the technique of "case" study is
    (*a*) the interview.
    (*b*) the summary and conclusion.
    (*c*) the school record.

76. The most important item in the history of a dull child is
    (*a*) the social history.
    (*b*) the educational history.
    (*c*) the medical history (including psychiatric).

77. The most significant aspect of a psychopath's emotional condition is
    (*a*) general tone.
    (*b*) emotional conflicts.
    (*c*) special emotional trends.

78. "Borderland" intelligence is frequently designated as
    (*a*) moronity.
    (*b*) amentia.
    (*c*) dementia.

79. Our knowledge is least dependable concerning
    (*a*) superior intelligence.
    (*b*) normality.
    (*c*) feeble-mindedness.

80. "Normality" is predominantly
    (*a*) a social concept.
    (*b*) an economic concept.
    (*c*) a psychologic concept.

81. The best descriptive term of most "borderland" cases is
    (*a*) "mentally peculiar"
    (*b*) "emotionally unstable."
    (*c*) "emotionally maladapted."

82. The psychopathic individual usually manifests
    (*a*) diseased conditions.
    (*b*) aberrancy of character.
    (*c*) character anomaly.

83. The "borderland" condition is most frequently
   (*a*) a mental abnormality induced by a physical condition.
   (*b*) solely a condition of anæmia.
   (*c*) a condition of emotional stress dislocating a normal personality.

84. "Psychoasthenia" is usually caused by
   (*a*) phobias.
   (*b*) family neuropathic taint.
   (*c*) insomnia.

85. To be adjudged "psychopathic," an individual must exhibit
   (*a*) hysteria and hallucinations.
   (*b*) marked emotional stress and emotional instability, noticeable from childhood.
   (*c*) marked emotional stress and striking lack of judgment of prolonged standing.

86. The rehabilitation of a psychopathic individual is a matter of
   (*a*) curing a condition.
   (*b*) changing an intractable personality.
   (*c*) altering the environment.

87. Most difficult "discipline cases" of the traditional school are recruited from among
   (*a*) the "eccentric."
   (*b*) the emotionally unstable.
   (*c*) the neurasthenic.

88. The most marked symptom of dementia præcox is
   (*a*) delusions of persecution.
   (*b*) disconnectedness in thinking.
   (*c*) a marked emotional indifference and apathy.

89. The feeble-minded are classified into (1) idiots, (2) imbeciles, (3) morons. These categories
   (*a*) are mutually exclusive.
   (*b*) constantly overlap.
   (*c*) are more descriptive than classificatory.

90. The criterion for diagnosing feeble-mindedness is
    (*a*) a rough standard.
    (*b*) a scientific "measuring rod."
    (*c*) a matter of expert judgment.

91. Feeble-mindedness
    (*a*) can be ameliorated.
    (*b*) can be cured.
    (*c*) can be retarded.

92. Feeble-mindedness is presumed in an individual who has a mental age of
    (*a*) twelve or lower.
    (*b*) ten or lower.
    (*c*) eight or lower.

93. The tendency is strong to place the responsibility for the amelioration of feeble-mindedness upon
    (*a*) eugenics.
    (*b*) the community.
    (*c*) the law.

94. The best type of education of the feeble-minded is conducted under
    (*a*) an honor system.
    (*b*) a monitorial system.
    (*c*) constant supervision.

95. The reported increase in feeble-mindedness is owing
    (*a*) to increases in the stresses imposed by modern civilization.
    (*b*) less to actual increase and more to improved method of diagnosing feeble-mindedness.
    (*c*) to increase in the number of feeble-minded immigrants.

96. Tredgold is the author of
    (*a*) *Technique of Psychoanalysis.*
    (*b*) *Mental Deficiency.*
    (*c*) *The Delinquent Girl.*

97. "Is exerting a tremendous influence on the testing movement":
    (*a*) Watson
    (*b*) Thorndike
    (*c*) Hollingworth

98. "Has modified profoundly the concept of intelligence":
    (*a*) Whipple
    (*b*) Monroe
    (*c*) Thurstone

99. "Is modifying the psychologic basis of learning":
    (*a*) Colvin
    (*b*) Kilpatrick
    (*c*) Koffka

100. "Has made the greatest contribution to the knowledge of the psychology of delinquency":
    (*a*) Kohs
    (*b*) McDougall
    (*c*) Burt

Number right.....Number unanswered.....
Number wrong.....Score = right $-\frac{1}{2}$ wrong =.....

SUMMARY OF STATISTICAL SYMBOLS, TERMS, AND FORMULÆ IN EDUCATIONAL MEASUREMENTS

Educational products are measurable. The measure is expressed in terms of units. These units are manifold and are distributed in increasing increments from very low to very high. The educational statistician seeks for certain information about this distribution. According to Truman L. Kelley,* there are six items of information requisite for the proper valuation of any distribution.

1. " The most important is what it is a distribution of."

2. " The second is the number of cases in the distribution or, as it is usually expressed, the population."

3. " The third is to know some measure of central tendency, some average." The most commonly employed measures of central tendency are (*a*) the mean (popularly known as the average), (*b*) the median, and (*c*) the mode.

4. " The fourth, to know some measure of the degree to which the measures scatter, or lie above and below the average, *i.e.*, to know a measure of dispersion or

* *Statistical Method*, Macmillan, 1923.

deviation from the average." (The term "average" is used generically.) The most important measures of dispersion are (*a*) the interquartile range (Q.), (*b*) the probable error (P.E.), and (*c*) the standard deviation (S.D.).

The fifth and sixth items are used by test makers who attempt an exhaustive statistical interpretation of the data under advisement. Most statistical manipulators proceed to the investigation of the degree of correlation subsisting between their test and an accepted criterion.

5. "The fifth, to know if the measures are symmetrically distributed with reference to the average or if there is a bunching of measures on one side of the average and a long trailing out of measures on the other side, *i.e.*, to know a measure of skewness."

6. "The sixth, to know if the measures are exceptionally densely grouped at the average, giving a high peak to the frequency polygon, or if the distribution is rather flat in the middle and contracted at the ends, thus tending toward a rectangular shape, or if they show a mean between those two conditions, as does a normal distribution" (which is bell-shaped).

The normal distribution is the one which begins at a point near zero, rises slowly to an apex indicated by the mean, median, or mode (they coincide in a perfect distribution), and slowly tapers off as it approaches zero among the largest units of the distribution.

Symbols Employed in Statistical Formulæ

N = Total number of measures (the total population)
X = Gross score (the summation from zero to the highest measure)
D = Deviation
x and y = Extent of deviation from the mean of two scores whose correlation is sought
M = Measure
f = Frequency (number of like measures)
fd (f times d) = Frequency deviation (total deviation)
M = Arithmetic mean (average)
Md or Mdn = Median
Mo = Mode
Q = Quartile deviation or semi-interquartile range
U.Q. = Upper quartile (100% to 75%)
L.Q. = Lower quartile (25% to 0)
A.D. = Average deviation (mean deviation from)
σ = Standard deviation from the mean, S.D.
Σ = Sum of scores indicated in the formula
P.E. = Probable error or median deviation
rho = Rank method coefficient of correlation
$P.E._{rho}$ = Probable error of rank method coefficient
r = Product-moment coefficient of correlation
$P.E._{r}$ = Probable error of product-moment coefficient

Basic Statistical Formulæ

1. (*a*) $M = \frac{\Sigma M}{N}$

M = Measures
N = Number of measures

(*b*) M (Shortened calculation) $= G + d - \frac{\Sigma \text{ (algebraic) deviations from } G}{N}$

G = Assumed or guessed mean
M = True (sought) mean value
Σ = Sum of
d = Deviation
N = Number of measures

2. Mdn (Median) $= \frac{N + 1}{2}$

3. (*a*) Standard Deviation from the Mean / S.D. $= \sqrt{\frac{\Sigma f d^2}{N}}$

or

S.D. = Square root of the arithmetic mean of the squares of deviations from the average of the distribution
d = Deviation from mean
f = Number of measures in each deviation
$\Sigma f d^2$ = Sum of the square of the product of f and d or the square of the total amount of deviation

Since the deviations are squared, the signs of deviation do not matter.

(*b*) S.D. (with guessed mean) $= \sqrt{\frac{\Sigma z^2}{N} - d^2}$

d = Algebraic sum of the individual variations from the guessed mean divided by N

N = Number of scores

Σz^2 = Sum of the squares of the deviations from the guessed mean

4. Q (Quartile deviation) $= \frac{Q_1 - Q_3}{2}$

Q_1 = Median of the upper 50% of frequencies arranged in a series

Q_3 = Median of lower 50%

5. P.E. $= .674495\,\sigma$

" It must be definitely kept in mind that the P.E. has no status or means of calculation independent of the standard error (S.D.); it is simply a measure of deviation .67449 times as large as the standard deviation." — Kelley

6. Amount of skewness $= \frac{\text{Mean} - \text{Mode}}{\sigma}$

" Because of the difficulty of calculating the mode with great exactness, and because the mode is usually about as far from the mean as is the median, Pearson's formula may for practical purposes be stated as follows:

$$\text{Skewness} = \frac{3\,(\text{Mean} - \text{Median})}{\sigma}$" — Kelley

7. Rank method coefficient of correlation

$$rho = 1 - \frac{6\,\Sigma d^2}{N(N-1)}$$

d = Deviation between the two ranks
d^2 = Sum of the squares of these deviations
N = Number of pupils or measures

8. Probable error of rank coefficient

$$\text{P.E.}_{rho} = .706\,\frac{1 - rho^2}{\sqrt{N}}$$

9. Product moment method coefficient of correlation

$$r = \frac{\Sigma x \cdot y}{\sqrt{\Sigma x^2}\ \sqrt{\Sigma y^2}}$$

" The coefficient of correlation is the quotient of the algebraic sum of the products of each x deviation and its corresponding y deviation, divided by the product of the square root of the sum of the squares of the x deviations multiplied by the square root of the sum of the squares of the y deviations." — Trabue

Suppose a correlation is sought between scores obtained by the same group in composition and grammar; then

x = Deviations of individual composition score from the mean for composition
y = Deviations of grammar scores from the mean for grammar
Σ = the algebraic sum

Probable error of product moment coefficient

$$\text{P.E.}_r = \frac{.67449(1 - r^2)}{\sqrt{N}}$$

r is considered more reliable than *rho*. Therefore a table has been worked out by which *rho* values are converted into r values.

Summary of Pedagogic Symbols and Ratios

Intelligence Ratios

M.A. = Mental age
C.A. = Chronological age
I.Q. = Intelligence quotient

Subject Ratios (worked out by McCall and Franzen)

R.A. = Reading age = Median reading score (on a certain test) for various ages
Ar.A = Arithmetic age
E.A. = Educational age = Average median score for many school subjects on certain specified tests

$$\text{E.Q.} = \text{Educational quotient} = \frac{\text{E.A.}}{\text{C.A.}}$$

$$\text{R.Q.} = \text{Reading quotient} = \frac{\text{R.A.}}{\text{C.A.}}$$

$$\text{Ar.Q.} = \text{Arithmetic quotient} = \frac{\text{Ar.A.}}{\text{C.A.}}$$

$$\text{A.Q. (McCall)} = \text{Accomplishment quotient} = \frac{\text{E.A.}}{\text{M.A.}}$$

$$\text{A.Q. (Franzen)} = \text{Accomplishment quotient} = \frac{\text{E.Q.}}{\text{I.Q.}}$$

$$\text{R.A.Q.} = \text{Reading accomplishment quotient} = \frac{\text{R.A.}}{\text{M.A.}}$$

$$\text{Ar.A.Q.} = \text{Arithmetic accomplishment quotient} = \frac{\text{Ar.A.}}{\text{M.A.}}$$

Bibliography

Alexander, G. — *School Statistics and Publicity*, Silver Burdett, 1919

Beately, E. E. — *I. A. Slide Rule*, Dobb and Co., Boston, 1922

Bowley, A. L. — *An Elementary Manual of Statistics*, MacDonald and Evans, London, 1910

Brown, W., and Thompson, G. H. — *Essentials of Mental Measurement*, Cambridge University Press, London, 1921

Buckingham, B. R. — *Research for Teachers*, Silver Burdett, 1926

Elderton, W. O. and E. M. — *Primer of Statistics*, A. and C. Black, London, 1910

Franzen, R. — "Attempts at Test Validation," *Journal of Educational Research*, September, 1922

Gregory, C. A. — *Elements of Educational Measurement*, Appleton, 1922

King, W. I. — *The Elements of Statistical Method*, Macmillan, 1912

McCall, W. A. — *How to Measure in Education*, Macmillan, 1922

Monroe, W. S. — *Measuring the Results of Teaching*, Houghton Mifflin, 1918

National Society for the Study of Education — *Twenty-first Year Book*, Part II, Public School Publishing Co., Bloomington, Illinois, 1922

Pearson, K. — *The Grammar of Science*, A. and C. Black, London, 1911

Pupils' Progress through the Grades, Board of Education, N. Y. C., Publication No. 19

Rugg, H. O. — *Statistical Methods Applied to Education*, Houghton Mifflin, 1917

Starch, D. — *Educational Measurements*, Macmillan, 1917

Thorndike, E. L. — *An Introduction to the Theory of Mental and Social Measurements*, Teachers College, Columbia University, 1916

West, C. J. — *Introduction to Mathematical Statistics*, R. G. Adams and Co., Columbus, Ohio, 1918

Whipple, G. M. — *Manual of Physical and Mental Tests*, Part 1, Warwick and York, 1920

INDEX

A

Accomplishment quotient (A.Q.), 20, 128
Accretive learning, 77, 97. *See* Learning
Accuracy of tests, 206
Achievement tests, 21; and dishonesty, 268; IER, 268. *See* Tests
Acquired equipment, 56
Adrenal bodies, 46
Arithmetic, errors in, 218
Arithmetic tests, 217, 224; and dishonesty, 268
Army tests, 20
Assimilative learning, 78, 98. *See* Learning
Attitudes, 66
Axis-cylinder, 41
Axone, 49
Ayres Handwriting Scale, 196, 226
Ayres Measuring Scale (spelling), 232

B

Badanes Observational Record, 246
Behavior, deceptive, measuring, 267
Behaviorists, 63. *See* Watson
Binet, A., 11, 13, 23, 137, 316
Binet revisions, 14
Binet Scale, 23, 150
Breuer, J., 336
Briggs English Form tests, 228
Buckingham Scale (arithmetic), 224
Burdick Apperception Tests, 279
Burgess Scale (silent reading), 230
Burks, B. S., 29
Burt revision, 143
Buswell and John Diagnostic Chart, 217

C

Cady, V. M., tests, 271
Cerebro-spinal system, 48
Chance hypothesis, 84
Character, measuring, 266
Charters Diagnostic Language Test, 229
Cheating, measuring, 267; aspersion of unmerited, 276. *See* Dishonesty
Chromosome, 41
Classroom, cheating in, 267
"Complex," 332
Composite estimate, 261
Conditioned emotional response, 61
Conditioned reflex, 58
Conditioning, 59
Configuration, 74
Configurative learning, 75, 82, 85, 96; levels of, 97
Control groups, 116
Conversion, 333
Coördination tests, 271
Correlation, between knowledge and conduct, 278
Criteria for devisal of tests, psychological, 287; administrative, 289
Cultural background and deception, 280

D

Dearborn Group Tests of Intelligence, 172
Deceit and deception, *see* Dishonesty
Defectives, 307, 329; characteristics of — bodily concomitants, 310, emotional concomitants, 311, mental, 312; criteria for evaluating — biological, 309, medical, 308, pedagogical, 308, psychological, 309, social-economic, 307
Deficiency, testing of, 314
Dementia praecox, 66
Dendrites, 41
"Detour," 85
"Detour-board," 84. *See* Configurative learning
Detroit First Grade Intelligence Test, 177
Detroit Kindergarten Test, 155
Diagnosis, 214, 221
Diagnostic chart, 217
Diagnostic tests, 216
Difficulty scales, 198
Digit Symbol Substitution Test, 269
Directive equipment, 83
Disarranged Sentence Test, 268
Dishonesty, 266; age and, 279; aspersion of unmerited, 276; correlation of influences and, 278; cultural background and, 280; distribution of cheating scores, 282; environment and, 277; health and, 279; heredity and, 277; intelligence and, 279; measuring methods, 267; norms in measuring, 274; race and religion and, 280; school conditions and, 280; sex and, 279; types of, 267
Distribution, valuation of, 383; of cheating scores, 282
Doctrine of specificness of conduct, attitudes, and motivation, 280
Doll, E. A., 28, 338
Double Testing Technique, 267; and dishonesty, 268
Downey, June, 253, 254
Ductless glands, 45
Duplicating Technique, 267
Dynamometer tests, 271

E

Educational age (E.A.), 127
Educational quotient (E.Q.), 127
Educational tests, 22. *See* Tests
Emotional bias, 220
Emotional stability, 221
Emotions, 53; classification of, 54
Endocrine glands, 45
Environment and dishonesty, 277
Epicritic movements, 58
Extraversion, 91

F

"Faculty" psychology, 2, 3, 109
Freeman Diagnostic Chart (handwriting), 227
Freud, S., 336
Freudians, 332

G

Galton, F., 8
Games, used in testing honesty, 267
Gauss, K. F., 8
Geography tests, 225
Gestalt and Gestaltists, 74, 155
Gifted Children, 313. *See* Superior children
Goddard, H., 15
Gray Oral Reading Tests, 231
Green, G. H., 240
Gregory-Spencer Geography Tests, 226
Group tests, 20. *See* Tests

H

Haggerty Intelligence Examination, 175
Haggerty Reading Examination, 231
Hahn-Lackey Scale (geography), 226
Handwriting scales, 226
Harland Test of Information (history), 228
Hartshorne, H., 266
Heredity and dishonesty, 277
Hillegas Nassau County Supplement (composition), 230
History tests, 228
Homogeneity of groups, 116
Honesty, 266; index of, 275; standard deviation, 274; value of tests of, 276. *See* Dishonesty, Tests
Hudelson Scale (composition), 229

I

Ideals of conduct and deceit, 280
Idiot, 7, 307
IER Achievement Tests, 268
Illinois Examination, 168
Imbecile, 307
Improbable achievement technique, 270
Information Test, 267
Innate equipment, 41, 51
Insight, 75
Instincts, 52
Institute of Educational Research, 268
Institute of Social and Religious Research, 266
Intelligence, Binet on, 106; Colvin on, 107; educational measure of, 127; Gestaltists on, 155; Stern on, 104; Terman on, 104; testing of, 107
Intelligence quotient (I.Q.), 19, 129, 130
Intelligence tests, *see* Tests
Introversion, 91
Itard, J. M., 6, 319

J

Janet, 335
Jung, C. G., 28, 91

K

Koffka, K., 100
Köhler, W., 72, 76, 78, 83, 99, 100
Kohs, S. C., 256
Kuhlmann Revision, 137

L

Laryngeal movement, 63
Lasurski, 244, 245
Learning by accretion, 77, 97; by assimilation, 77, 198; by gradation, 78, 98; by redefinition, 80, 99
Learning process, 68
Lewis Special-Type Scales (composition), 229
Lie index, 275
Long-distance memory, 260
Lying, measuring, 267, 273. *See* Dishonesty

M

McDougall, 53
"Mass tests," 109
Mateer, F., 321
May, M. A., and Hartshorne, H., 263
Mayo, A. D., 331
Measuring dishonesty, median and percentile scores, 277; methods, 267; norms, 276; percentage of

error, 276; use of Sims Score Card, 279
Median, 386
Memory, quality of, 260
Memory span, 119
Mental age (M.A.), 129
Mental defective, 332. *See* Defective
Methods of measuring dishonesty, 267. *See* Dishonesty
Monroe Diagnostic Tests (arithmetic), 224
Monroe Silent Reading Tests, 231
Monroe Standard Reasoning Test (arithmetic), 231
Monroe Timed Sentence Test, 232
Montessori, M., 322
Multiple testing technique, 267
Mutism, 330
Myers Mental Measure, 170

N

National Intelligence Test, 165, 269
Neural coördination, 56
Neurone, 41, 48
Neurotic, 329, 331. *See* Psychopathy
"New type" tests, 212
Non-intelligence traits, 238
Normal curve of distribution, 9
Norms, in measuring dishonesty, their standardization and interpretation, 274
Nucleus, 41

O

Observational Record, 246
Ogden, R. M., 241
Organs, conducting, 48; of response, 44; visceral, 44
Otis Group Intelligence Scale, 162
Otis Self-Administering Tests, 180

P

Pattern, 78. *See* Configuration
Percentile rank, 201
Personal Data Sheet, 273
Personality, analysis of, 38; biological interpretation of, 27; building, 48; classifications of, 88, 89; definitions of, 88, 89; traits, 238; significance of, 241
Pintner-Paterson Performance Scale, 16
Pressey First Grade Reading Vocabulary Test, 232
Probable error (P.E.), 197, 200; locations, 201, 202
Prognosis, *see* Diagnosis
Protoplasm, 41
Psychograph, 28
Psychopathy, causes of, 330; means of determining, 338
Pupil Data Sheet, 273
Purposiveness, 70, 73
Puzzle performance tests, 270

Q

Quetelet, L. A. J., 8
Quotient, intelligence, *see* I.Q.

R

Rationalization, 334
Reasoning test, 154
Reflex arc, 49
Reliability of tests, 205, 206
Rice, J. M., 21
Rivers, R. H., 56
Rossolimo, 28

S

Scale, definition of, 122; qualities of, 114, 195. *See* Tests *and* Tests and scales

Scattered scores, 116
School conditions and dishonesty, 280
Sequin, E., 320
Sense organs, 43
Sentence Completion Test, 268
Skewness, 387
Span of attention, 118
Spearman, C., 17
Specificness of conduct, attitudes, and motivation, 280
Spelling and dishonesty, 268
Standard deviation (S.D.) 199, 274, 386
Stanford Revision, 137
Statistics, educational, 8, 17, 19, 115, 129, 200
Stealing, measuring, 267, 272
Steinach, 48
Stevens Arithmetic Reading Scale, 225
Stimulus, unique, 62
Studies in Deceit, 266; doctrines of, 280; evaluation of technique of, 282; practical value of tests in, 276
Superior child, 349
Superior children, education of, 359; studies of, 350; method of grouping, 354
Sympathetic system, 49
Synthetic tests, 109

T

Teachers' estimates, 258
Technique of measuring honesty, etc., 267; norms for, 276; evaluation of, 282
Temperament, 89
Terman, L. M., 15, 29, 137
Tests, administrative qualities of, 119; allocation of, 148; alternative, 26; construction of, 113, 192, 287; direction of improvement, 292; material, 108, 113, 192, 287; pedagogical qualities of, 119; statistical qualities of, 114. *See* Criteria
Tests (*see* Tests and scales; *see also* Measuring dishonesty):

NON-INTELLIGENCE TRAITS

Allport Active-Passive Study, 297
Burdick Apperception Test, 279
Cady Moral Judgment Test, 271, 299
Downey Group-Will Temperament Test, 254
Fernald Achievement Capacity Test, 298
IER, 268
Kohs Moral Discrimination, 256
Myerson Personality Tests, 299
Pressey Mental Survey, 298
Pressey X–O Tests, 298
Sims Score Card, 279
Thorndike-McCall Reading Scale, 268
Voelker Honesty Test, 271, 299
Voelker Reliability Test, 299
Voelker Tracing Opposite Test, 298
Watson Survey of Public Opinion, 297
Watts Association Test, 339
Woodworth Personal Data Sheet, 297

Tests and scales:

ACHIEVEMENT

Ayres Handwriting Scale, 196, 226
Ayres Measuring Scale (spelling), 232
Briggs English Form Tests, 228
Buckingham Scale (arithmetic), 224

Burgess Scale (silent reading), 230
Buswell-John Diagnostic Chart, 217
Charters Diagnostic Language Test, 229
Courtis Standard Practice Test (handwriting), 227
Courtis Standard Research Test (arithmetic), 224
Courtis Standard Supervisory Test (geography), 225
Freeman Diagnostic Chart (handwriting), 227
Gray Oral Reading Test, 231
Gregory-Spencer Geography Tests, 226
Haggerty Reading Examination, 231
Hahn-Lackey Scale (geography), 226
Harlan Test of Information (history), 228
Hillegas Nassau County Supplement (composition), 230
Hudelson Scale (composition), 229
IER Achievement Scales, 268
Lewis Special-Type Scales (composition), 229
Monroe Diagnostic Tests (arithmetic), 224
Monroe Silent Reading Tests, 231
Monroe Standardized Reasoning Tests (arithmetic), 231
Monroe Timed Sentence Test, 232
Pressey First Grade Reading Vocabulary Test, 232
Stevenson Arithmetic Reading Test, 225
Thorndike Handwriting Scale, 227
Van Wagenen History Scales, 228
Woody Arithmetic Scales, 225
Woody-McCall Mixed Fundamentals, 225

INTELLIGENCE

Burt Reasoning Test, 154
Burt Revision, 143
Dearborn Group Tests of Intelligence, 172
Detroit First Grade Intelligence Test, 177
Detroit Kindergarten Test, 155
Haggerty Intelligence Examination, 175
Illinois Examination, 168
Kuhlmann Revision, 141
Myers Mental Measure, 170
National Intelligence Examination, 165
Otis Group Intelligence Scale, 162
Otis Self-Administering Tests, 165
Porteus Maze, 152
Stanford Revision, 137
Trabue Mentimeters, 183
Testing material, 24
Testing movement, biological qualitative phase, 23; psychological qualitative phase, 6; psychological quantitative phase, 8; sociological qualitative phase, 4
Testing program, 295
Thorndike, E. L., 21, 68, 83, 267
Thorndike Handwriting Scale, 227
Thorndike-McCall Reading Scale, 268
Thyroids, 46
Thyroxin, 45
Tissue, connective, 42; epithelial, 42; nervous, 42

Tredgold, A. F., 318
Trial and error learning, 68, 79. *See* Learning
Triple testing technique, 269
T scores, 204, 275

V

Validity of tests, 205, 206
Van Wagenen History Scales, 228
Verbalism, 330
Voelker Honesty Test, 271, 299
Voelker Reliability Test, 299
Voelker Tracing and Opposites Test, 298
Volitional Perseveration, 255
Voronoff, 48

W

Watson, J. B., 29, 54
Watts, F., 339
Will-Temperament Test, 254
Woody Arithmetic Scale, 225
Woody-McCall Mixed Fundamentals, 225
Word Knowledge Completion Test, 268

Y

Young, K., 17, 26

Z

Ziehen, 330

BF 431 .L4 32,020

Levine, Albert Julius.

Testing intelligence and achievement,

BF 431 .L4 32,020

Levine, Albert Julius.

Testing intelligence and achievement,

UNITED THEOLOGICAL SEMINARY

United Theological Seminary Library
MAIN BF431.L4
Levine, Alber Testing intelligence and a
3 0645 0003 2045 3